

414 Rec'd PCT/PTO 02 NOV 2000

PCT

| | | |
|--|--|---|
| FORM PTO-1390 U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371 | | ATTORNEY'S DOCKET NUMBER PF-0509 US |
| INTERNATIONAL APPLICATION NO. PCT/US99/09935 | | U.S. APPLICATION NO. (If known, see 37 CFR 1.5) To Be Filed 091 674 743 |
| TITLE OF INVENTION HUMAN TRANSCRIPTIONAL REGULATOR MOLECULES | | PRIORITY DATE CLAIMED 05 May 1998 |
| APPLICANT(S) FOR DO/EO/US INCYTE PHARMACEUTICALS, INC.; HILLMAN, Jennifer L.; BANDMAN, Olga; LAL, Preeti; YUE, Henry; REDDY, Roopa; TANG, Y. Tom; GERSTIN, Edward H.; PATTERSON, Chandra; BAUGHN, Mariah R.; AZIMZAI, Yalda; LU, Duyng Aina M. | | |
| <p>Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:</p> <ol style="list-style-type: none"> 1. <input checked="" type="checkbox"/> This is the FIRST submission of items concerning a filing under 35 U.S.C. 371 2. <input type="checkbox"/> This is a SECOND or SUBSEQUENT submission of items concerning a filing under 35 U.S.C. 371 3. <input type="checkbox"/> This is an express request to promptly begin national examination procedures (35 U.S.C. 371 (f)). 4. <input type="checkbox"/> The US has been elected by the expiration of 19 months from the priority date (PCT Article 31). 5. <input checked="" type="checkbox"/> A copy of the International Application as filed (35 U.S.C. 371(c)(2)) <ol style="list-style-type: none"> a. <input type="checkbox"/> is attached hereto (required only if not communicated by the International Bureau) b. <input type="checkbox"/> has been communicated by the International Bureau. c. <input checked="" type="checkbox"/> is not required, as the application was filed in the United States Receiving Office (RO/US). 6. <input type="checkbox"/> An English language translation of the International Application as filed (35 U.S.C. 371(c)(2)). 7. <input checked="" type="checkbox"/> Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371(c)(3)) <ol style="list-style-type: none"> a. <input type="checkbox"/> are attached hereto (required only if not communicated by the International Bureau) b. <input type="checkbox"/> have been communicated by the International Bureau. c. <input type="checkbox"/> have not been made, however, the time limit for making such amendments has NOT expired. d. <input checked="" type="checkbox"/> have not been made and will not be made. 8. <input type="checkbox"/> An English language translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3)). 9. <input type="checkbox"/> An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4)). 10. <input type="checkbox"/> An English language translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371(c)(5)) | | |
| Items 11 to 16 below concern document(s) or information included: | | |



09/674743
526 Rec'd PCT/TTO 02NOV2000

11. An Information Disclosure Statement under 37 CFR 1.97 and 1.98.
12. An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.27 and 3.31 is included.
13. A FIRST preliminary amendment.
 A SECOND or SUBSEQUENT preliminary amendment.
14. A substitute specification.
15. A change of power of attorney and/or address letter.
16. Other items or information:

- 1) Transmittal Letter (3 pp, in duplicate)
2) Return Postcard
3) Express Mail Label No.: EL 579 975 901

| | | | |
|--|--------------|--|--|
| U.S. APPLICATION NO. (if known, see 37 CFR 1.55) To Be Assigned | | INTERNATIONAL APPLICATION NO PCT/US99/09935 | ATTORNEY'S DOCKET NUMBER PF-0509 US |
| 17. <input type="checkbox"/> The following fees are submitted: BASIC NATIONAL FEE (37 CFR 1.492(a)(1)-(5)) Neither international preliminary examination fee (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2)) paid to USPTO and International Search Report not prepared by the EPO or JPO.\$1000.00 International preliminary examination fee (37 CFR 1.482) not paid to USPTO but International Search Report prepared by the EPO or JPO.\$860.00 International preliminary examination fee (37 CFR 1.482) not paid to USPTO but international search fee (37 CFR 1.445(a)(2)) paid to USPTO.\$710.00 <input type="checkbox"/> International preliminary examination fee paid to USPTO (37 CFR 1.482) but all claims did not satisfy provisions of PCT Article 33(1)(4).\$690.00 International preliminary examination fee paid to USPTO (37 CFR 1.482) and all claims satisfied provisions of PCT Article 33(1)(4).\$100.00 | | | |
| ENTER APPROPRIATE BASIC FEE AMOUNT = \$690.00 | | | |
| Surcharge of \$130.00 for furnishing the oath or declaration later than <input type="checkbox"/> 20 <input type="checkbox"/> 30 months from the earliest claimed priority date (37 CFR 1.492(e)). \$ | | | |
| CLAIMS | NUMBER FILED | NUMBER EXTRA | RATE |
| Total Claims | - 20 - | | X \$ 18.00 \$ |
| Independent Claims | - 3 - | | X \$ 80.00 \$ |
| MULTIPLE DEPENDENT CLAIM(S) (if applicable) | | + \$270.00 \$ | |
| TOTAL OF ABOVE CALCULATIONS = \$690.00 | | | |
| <input type="checkbox"/> Applicant claims small entity status. See 37 CFR 1.27. The fees indicated above are reduced by 1/2. \$ | | | |
| SUBTOTAL = \$690.00 | | | |
| Processing fee of \$130.00 for furnishing the English translation later than <input type="checkbox"/> 20 <input type="checkbox"/> 30 months from the earliest claimed priority date (37 CFR 1.492(i)). \$ | | | |

09/674743

526 Rec'd PCT/PTO 02 NOV 2000

| | | | |
|---|--|-----------------------|----|
| TOTAL NATIONAL FEE = | | \$690.00 | |
| Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by the appropriate cover sheet (37 CFR 3.28, 3.31). \$40.00 per property + | | \$ | |
| TOTAL FEES ENCLOSED = | | \$690.00 | |
| | | Amount to be Refunded | \$ |
| | | Charged | \$ |
| a. <input type="checkbox"/> A check in the amount of \$_____ to cover the above fees is enclosed. | | | |
| b. <input checked="" type="checkbox"/> Please charge my Deposit Account No. 09-0108 in the amount of \$ 690.00 to cover the above fees. | | | |
| c. <input checked="" type="checkbox"/> The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. 09-0108. A duplicate copy of this sheet is enclosed. | | | |
| NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b)) must be filed and granted to restore the application to pending status. | | | |
| SEND ALL CORRESPONDENCE TO: | | | |
| INCYTE GENOMICS, INC. 3160 Porter Drive Palo Alto, CA 94304 | | | |
|  SIGNATURE <u>2 November 2000</u> <hr/> NAME: Diana Hamlet-Cox | | | |
| <hr/> REGISTRATION NUMBER: 33,302 | | | |

PROMPT DATE 02 NOV 2000

HUMAN TRANSCRIPTIONAL REGULATOR MOLECULES

5

TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of human transcriptional regulator molecules and to the use of these sequences in the diagnosis, treatment, and prevention of cell proliferative and immune disorders.

10

BACKGROUND OF THE INVENTION

Differential control of gene expression is essential to the growth and development of all multicellular organisms. Although gene expression can be controlled at many steps along the path from DNA to protein, the major control point for most genes is at the initiation of transcription. This critical step is regulated both positively and negatively by a combination of general and tissue specific transcription factors, the majority of which function to regulate transcription of one or more target genes.

Mutations in transcription factors (TFs) contribute to oncogenesis. This is probably due to the role of transcription factors on the expression of genes involved in cell proliferation. For example, mutations in transcription factors encoded by proto-oncogenes, such as Fos, Jun, Myc, Rel, and Spi-1, may be oncogenic due to increased stimulation of cell proliferation. Conversely, mutations in transcription factors encoded by tumor suppressor genes, such as p53, RB1, and WT1, may be oncogenic due to decreased inhibition of cell proliferation. (Latchman, D. (1995) Gene Regulation: A Eukaryotic Perspective, Chapman and Hall, London, UK, pp 242-255.)

Many transcription factors are modular proteins that contain separate domains for DNA binding and transcriptional regulation. The DNA binding domain interacts with specific DNA sequences (control elements) near to or within the promoter region of the gene. This interaction brings the regulatory domain of the TF into a position where it can interact with other proteins to stimulate or repress transcription. Many TFs require dimerization or multimerization to be fully functional. Five different types of transcription factors have been described based on five well characterized structural motifs. These five types are the helix-turn-helix, zinc finger, leucine zipper, and helix-loop-helix (HLH) proteins and the steroid-hormone receptors.

The helix-turn-helix motif consists of two α helices held at a fixed angle. The two helices are connected by a short chain of amino acids, which represents the "turn". The more carboxyl-terminal helix is called the recognition helix and fits into the major groove of the DNA double helix. The recognition helix, whose amino acid side chains differ from protein to protein, plays an

important role in recognizing the specific DNA sequence to which the protein binds. All of the helix-turn-helix proteins bind DNA as dimers in which the two copies of the recognition helix are separated by exactly one turn of the DNA helix. Homeodomain proteins are a special class of helix-turn-helix protein. The homeodomain is folded into three α helices which are packed tightly together by hydrophobic interactions. Helices two and three closely resemble the helix-turn-helix motif, with the third helix acting as the recognition helix. Proteins containing homeodomain motifs often function as developmental switches.

The zinc finger motif consists of an α helix and antiparallel β sheet held together by a zinc atom. The zinc finger motif is usually repeated in a tandem array within a protein, such that the α helix of each zinc finger in the protein makes contact with the major groove of the DNA double helix. This repeated contact between the protein and the DNA produces a strong and specific DNA-protein interaction. The strength and specificity of the interaction can be regulated by the number of zinc finger motifs within the protein.

The leucine zipper motif consists of a single α helix which is involved in both protein dimerization and DNA binding. Two proteins containing leucine zippers can dimerize by interactions between hydrophobic amino acid residues, commonly leucines, that extend from one side of their respective α helices. In this way, the α helices of each protein monomer dimerize to form a short coiled-coil. Just beyond this coiled-coil, the two α helices separate to form a Y-shaped structure which contacts the major groove of the DNA. Leucine zipper proteins may form homodimers, in which the two protein monomers are identical, or heterodimers, in which the two protein monomers are different. The specificity of DNA binding depends on the dimer formed, since each protein monomer has distinct DNA-binding specificities.

The helix-loop-helix (HLH) motif consists of a short α helix connected by a loop to a second, longer α helix. The flexible loop allows the two helices to fold back and pack together. As with the leucine zipper, the HLH motif is involved in both protein dimerization and DNA-binding. The dimers can be homodimers or heterodimers, thus increasing the repertoire of DNA-binding sites to which HLH proteins can bind.

The steroid-hormone receptors contain a motif composed of two perpendicular α helices. In the absence of ligand the steroid-hormone receptors assume a conformation which sequesters the α helices. Binding of ligand, commonly steroid hormones, thyroid hormones, retinoids, or vitamin D, to the receptor causes a conformational change which exposes the α helices. The first α helix contains about seventy residues and includes eight conserved cysteines. This helix fits into the major groove of the DNA double helix and enables DNA-receptor binding. The second α helix provides for protein dimerization. As with leucine zipper and HLH proteins, both homodimers and heterodimers may be formed by steroid-hormone receptors.

- Hundreds of regulatory proteins from a wide variety of organisms have been identified. Most of these proteins have at least one of the common structural motifs described. However, several important regulatory proteins, including the p53 tumor suppressor, have a unique structure not shared with other known regulatory molecules. (Faisst, S. and S. Meyer (1992) Nucl. Acids Res. 20:3-26.) Moreover, other domains of the regulatory proteins often form crucial contacts with the DNA, thereby affecting binding specificity. Accessory proteins can also provide important interactions which may convert a particular regulatory protein from an activator to a repressor, from a repressor to an activator, or it may prevent DNA binding by the regulatory protein completely.
- 10 The discovery of new human transcriptional regulator molecules and the polynucleotides encoding them satisfies a need in the art by providing new compositions which are useful in the diagnosis, prevention, and treatment of cell proliferative and immune disorders.

SUMMARY OF THE INVENTION

- 15 The invention features substantially purified polypeptides, human transcriptional regulator molecules, referred to collectively as "HTRM". In one aspect, the invention provides a substantially purified polypeptide comprising an amino acid sequence selected from the group consisting of
- 20 SEQ ID NO:1-65, and fragments thereof.
- The invention further provides a substantially purified variant having at least 90% amino acid identity to at least one of the amino acid sequences selected from the group consisting of SEQ ID NO:1-65, and fragments thereof. The invention also provides an isolated and purified polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of
- 25 SEQ ID NO:1-65, and fragments thereof. The invention also includes an isolated and purified polynucleotide variant having at least 70% polynucleotide sequence identity to the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of
- 30 of
- SEQ ID NO:1-65, and fragments thereof.
- Additionally, the invention provides an isolated and purified polynucleotide which hybridizes under stringent conditions to the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-65, and fragments thereof. The invention also provides an isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide encoding the polypeptide comprising the amino

acid sequence selected from the group consisting of SEQ ID NO:1-65, and fragments thereof.

The invention also provides an isolated and purified polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:66-130, and fragments thereof. The invention further provides an isolated and purified polynucleotide variant having at least 70% polynucleotide sequence identity to the polynucleotide sequence selected from the group consisting of SEQ ID NO:66-130, and fragments thereof. The invention also provides an isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:66-130, and fragments thereof.

10 The invention also provides a method for detecting a polynucleotide in a sample containing nucleic acids, the method comprising the steps of (a) hybridizing the complement of the polynucleotide sequence to at least one of the polynucleotides of the sample, thereby forming a hybridization complex; and (b) detecting the hybridization complex, wherein the presence of the hybridization complex correlates with the presence of a polynucleotide in the sample. In one aspect, the method further comprises amplifying the polynucleotide prior to hybridization.

15 The invention further provides an expression vector containing at least a fragment of the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-65, and fragments thereof. In another aspect, the expression vector is contained within a host cell.

20 The invention also provides a method for producing a polypeptide, the method comprising the steps of: (a) culturing the host cell containing an expression vector containing at least a fragment of a polynucleotide under conditions suitable for the expression of the polypeptide; and (b) recovering the polypeptide from the host cell culture.

25 The invention also provides a pharmaceutical composition comprising a substantially purified polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NO:1-65, and fragments thereof, in conjunction with a suitable pharmaceutical carrier.

The invention further includes a purified antibody which binds to a polypeptide selected from the group consisting of SEQ ID NO:1-65, and fragments thereof. The invention also provides a purified agonist and a purified antagonist to the polypeptide.

30 The invention also provides a method for treating or preventing a disorder of cell proliferation associated with decreased expression or activity of HTRM, the method comprising administering to a subject in need of such treatment an effective amount of a pharmaceutical composition comprising a substantially purified polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NO:1-65, and fragments thereof, in conjunction 35 with a suitable pharmaceutical carrier.

The invention also provides a method for treating or preventing a disorder of cell proliferation associated with increased expression or activity of HTRM, the method comprising administering to a subject in need of such treatment an effective amount of an antagonist of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-5 65, and fragments thereof.

BRIEF DESCRIPTION OF THE TABLES

Table 1 shows nucleotide and polypeptide sequence identification numbers (SEQ ID NO), clone identification numbers (clone ID), cDNA libraries, and cDNA fragments used to assemble 10 full-length sequences encoding HTRM.

Table 2 shows features of each polypeptide sequence including potential motifs, homologous sequences, and methods and algorithms used for identification of HTRM.

Table 3 shows the tissue-specific expression patterns of each nucleic acid sequence as determined by northern analysis, diseases, disorders, or conditions associated with these tissues, 15 and the vector into which each cDNA was cloned.

Table 4 describes the tissues used to construct the cDNA libraries from which Incyte cDNA clones encoding HTRM were isolated.

Table 5 shows the programs, their descriptions, references, and threshold parameters used to analyze HTRM.

20

DESCRIPTION OF THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular machines, materials and methods described, as these may vary. It is also to be understood that the terminology used herein is for the 25 purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an 30 antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described 35 herein can be used to practice or test the present invention, the preferred machines, materials and

methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of 5 prior invention.

DEFINITIONS

"HTRM" refers to the amino acid sequences of substantially purified HTRM obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and preferably the human species, from any source, whether natural, synthetic, 10 semi-synthetic, or recombinant.

The term "agonist" refers to a molecule which, when bound to HTRM, increases or prolongs the duration of the effect of HTRM. Agonists may include proteins, nucleic acids, carbohydrates, or any other molecules which bind to and modulate the effect of HTRM.

An "allelic variant" is an alternative form of the gene encoding HTRM. Allelic variants 15 may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. Any given natural or recombinant gene may have none, one, or many allelic forms. Common mutational changes which give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination 20 with the others, one or more times in a given sequence.

"Altered" nucleic acid sequences encoding HTRM include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polynucleotide the same as HTRM or a polypeptide with at least one functional characteristic of HTRM. Included within this definition are polymorphisms which may or may not be readily detectable using a particular 25 oligonucleotide probe of the polynucleotide encoding HTRM, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding HTRM. The encoded protein may also be "altered," and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent HTRM. Deliberate amino acid substitutions may be made 30 on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of HTRM is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, positively charged amino acids may include lysine and arginine, and amino acids with uncharged polar head groups having similar hydrophilicity values may include leucine, isoleucine, 35 and valine; glycine and alanine; asparagine and glutamine; serine and threonine; and

phenylalanine and tyrosine.

- The terms "amino acid" or "amino acid sequence" refer to an oligopeptide, peptide, polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. In this context, "fragments," "immunogenic fragments," or "antigenic fragments" refer to fragments of HTRM which are preferably at least 5 to about 15 amino acids in length, most preferably at least 14 amino acids, and which retain some biological activity or immunological activity of HTRM. Where "amino acid sequence" is recited to refer to an amino acid sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated 10 with the recited protein molecule.

"Amplification" relates to the production of additional copies of a nucleic acid sequence. Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known in the art.

- The term "antagonist" refers to a molecule which, when bound to HTRM, decreases the 15 amount or the duration of the effect of the biological or immunological activity of HTRM. Antagonists may include proteins, nucleic acids, carbohydrates, antibodies, or any other molecules which decrease the effect of HTRM.

- The term "antibody" refers to intact molecules as well as to fragments thereof, such as Fab, F(ab')₂, and Fv fragments, which are capable of binding the epitopic determinant. Antibodies 20 that bind HTRM polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, 25 and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

- The term "antigenic determinant" refers to that fragment of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (given regions or three-dimensional structures on 30 the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

- The term "antisense" refers to any composition containing a nucleic acid sequence which is complementary to the "sense" strand of a specific nucleic acid sequence. Antisense molecules may be produced by any method including synthesis or transcription. Once introduced into a cell, 35 the complementary nucleotides combine with natural sequences produced by the cell to form

duplexes and to block either transcription or translation. The designation "negative" can refer to the antisense strand, and the designation "positive" can refer to the sense strand.

The term "biologically active," refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" refers to the capability of the natural, recombinant, or synthetic HTRM, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The terms "complementary" or "complementarity" refer to the natural binding of polynucleotides by base pairing. For example, the sequence "5' A-G-T 3'" bonds to the 10 complementary sequence "3' T-C-A 5'." Complementarity between two single-stranded molecules may be "partial," such that only some of the nucleic acids bind, or it may be "complete," such that total complementarity exists between the single stranded molecules. The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of the hybridization between the nucleic acid strands. This is of particular importance in 15 amplification reactions, which depend upon binding between nucleic acids strands, and in the design and use of peptide nucleic acid (PNA) molecules.

A "composition comprising a given polynucleotide sequence" or a "composition comprising a given amino acid sequence" refer broadly to any composition containing the given polynucleotide or amino acid sequence. The composition may comprise a dry formulation or an 20 aqueous solution. Compositions comprising polynucleotide sequences encoding HTRM or fragments of HTRM may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, 25 dry milk, salmon sperm DNA, etc.).

"Consensus sequence" refers to a nucleic acid sequence which has been resequenced to resolve uncalled bases, extended using XL-PCR kit (Perkin-Elmer, Norwalk CT) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from the overlapping sequences of more than one Incyte Clone using a computer program for fragment assembly, such as the 30 GELVIEW Fragment Assembly system (GCG, Madison WI). Some sequences have been both extended and assembled to produce the consensus sequence.

The term "correlates with expression of a polynucleotide" indicates that the detection of the presence of nucleic acids, the same or related to a nucleic acid sequence encoding HTRM, by northern analysis is indicative of the presence of nucleic acids encoding HTRM in a sample, and 35 thereby correlates with expression of the transcript from the polynucleotide encoding HTRM.

A "deletion" refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

The term "derivative" refers to the chemical modification of a polypeptide sequence, or a polynucleotide sequence. Chemical modifications of a polynucleotide sequence can include, for example, replacement of hydrogen by an alkyl, acyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

10 The term "similarity" refers to a degree of complementarity. There may be partial similarity or complete similarity. The word "identity" may substitute for the word "similarity." A partially complementary sequence that at least partially inhibits an identical sequence from hybridizing to a target nucleic acid is referred to as "substantially similar." The inhibition of hybridization of the completely complementary sequence to the target sequence may be examined 15 using a hybridization assay (Southern or northern blot, solution hybridization, and the like) under conditions of reduced stringency. A substantially similar sequence or hybridization probe will compete for and inhibit the binding of a completely similar (identical) sequence to the target sequence under conditions of reduced stringency. This is not to say that conditions of reduced stringency are such that non-specific binding is permitted, as reduced stringency conditions 20 require that the binding of two sequences to one another be a specific (i.e., a selective) interaction. The absence of non-specific binding may be tested by the use of a second target sequence which lacks even a partial degree of complementarity (e.g., less than about 30% similarity or identity). In the absence of non-specific binding, the substantially similar sequence or probe will not hybridize to the second non-complementary target sequence.

25 The phrases "percent identity" or "% identity" refer to the percentage of sequence similarity found in a comparison of two or more amino acid or nucleic acid sequences. Percent identity can be determined electronically, e.g., by using the MEGALIGN program (DNASTAR, Madison WI) which creates alignments between two or more sequences according to methods selected by the user, e.g., the clustal method. (See, e.g., Higgins, D.G. and P.M. Sharp (1988) 30 Gene 73:237-244.) The clustal algorithm groups sequences into clusters by examining the distances between all pairs. The clusters are aligned pairwise and then in groups. The percentage similarity between two amino acid sequences, e.g., sequence A and sequence B, is calculated by dividing the length of sequence A, minus the number of gap residues in sequence A, minus the number of gap residues in sequence B, into the sum of the residue matches between sequence A 35 and sequence B, times one hundred. Gaps of low or no similarity between the two amino acid

sequences are not included in determining percentage similarity. Percent identity between nucleic acid sequences can also be counted or calculated by other methods known in the art, e.g., the Jotun Hein method. (See, e.g., Hein, J. (1990) Methods Enzymol. 183:626-645.) Identity between sequences can also be determined by other methods known in the art, e.g., by varying

5 hybridization conditions.

"Human artificial chromosomes" (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size, and which contain all of the elements required for stable mitotic chromosome segregation and maintenance.

The term "humanized antibody" refers to antibody molecules in which the amino acid 10 sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

"Hybridization" refers to any process by which a strand of nucleic acid binds with a complementary strand through base pairing.

The term "hybridization complex" refers to a complex formed between two nucleic acid 15 sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C₆t or R₆t analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

20 The words "insertion" or "addition" refer to changes in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively, to the sequence found in the naturally occurring molecule.

"Immune response" can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by 25 expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect cellular and systemic defense systems.

The term "microarray" refers to an arrangement of distinct polynucleotides on a substrate.

The terms "element" or "array element" in a microarray context, refer to hybridizable polynucleotides arranged on the surface of a substrate.

30 The term "modulate" refers to a change in the activity of HTRM. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of HTRM.

The phrases "nucleic acid" or "nucleic acid sequence" refer to a nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or 35 RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may

represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material. In this context, "fragments" refers to those nucleic acid sequences which, when translated, would produce polypeptides retaining some functional characteristic, e.g., antigenicity, or structural domain characteristic, e.g., ATP-binding site, of the full-length

5 polypeptide.

The terms "operably associated" or "operably linked" refer to functionally related nucleic acid sequences. A promoter is operably associated or operably linked with a coding sequence if the promoter controls the translation of the encoded polypeptide. While operably associated or operably linked nucleic acid sequences can be contiguous and in the same reading frame, certain

10 genetic elements, e.g., repressor genes, are not contiguously linked to the sequence encoding the polypeptide but still bind to operator sequences that control expression of the polypeptide.

The term "oligonucleotide" refers to a nucleic acid sequence of at least about 6 nucleotides to 60 nucleotides, preferably about 15 to 30 nucleotides, and most preferably about 20 to 25 nucleotides, which can be used in PCR amplification or in a hybridization assay or

15 microarray. "Oligonucleotide" is substantially equivalent to the terms "amplimer," "primer," "oligomer," and "probe," as these terms are commonly defined in the art.

"Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition.

20 PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

The term "sample" is used in its broadest sense. A sample suspected of containing nucleic acids encoding HTRM, or fragments thereof, or HTRM itself, may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic

25 DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

The terms "specific binding" or "specifically binding" refer to that interaction between a protein or peptide and an agonist, an antibody, or an antagonist. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope,

recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the

30 presence of a polypeptide containing the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

The term "stringent conditions" refers to conditions which permit hybridization between polynucleotides and the claimed polynucleotides. Stringent conditions can be defined by salt

35 concentration, the concentration of organic solvent, e.g., formamide, temperature, and other

conditions well known in the art. In particular, stringency can be increased by reducing the concentration of salt, increasing the concentration of formamide, or raising the hybridization temperature.

- The term "substantially purified" refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least about 60% free, preferably about 75% free, and most preferably about 90% free from other components with which they are naturally associated.

A "substitution" refers to the replacement of one or more amino acids or nucleotides by different amino acids or nucleotides, respectively.

- "Substrate" refers to any suitable rigid or semi-rigid support including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

- "Transformation" describes a process by which exogenous DNA enters and changes a recipient cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to, viral infection, electroporation, heat shock, lipofection, and particle bombardment.
- The term "transformed" cells includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

- A "variant" of HTRM polypeptides refers to an amino acid sequence that is altered by one or more amino acid residues. The variant may have "conservative" changes, wherein a substituted amino acid has similar structural or chemical properties (e.g., replacement of leucine with isoleucine). More rarely, a variant may have "nonconservative" changes (e.g., replacement of glycine with tryptophan). Analogous minor variations may also include amino acid deletions or insertions, or both. Guidance in determining which amino acid residues may be substituted, inserted, or deleted without abolishing biological or immunological activity may be found using computer programs well known in the art, for example, LASERGENE software (DNASTAR).

- The term "variant," when used in the context of a polynucleotide sequence, may encompass a polynucleotide sequence related to HTRM. This definition may also include, for example, "allelic" (as defined above), "splice," "species," or "polymorphic" variants. A splice variant may have significant identity to a reference molecule, but will generally have a greater or

lesser number of polynucleotides due to alternate splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or an absence of domains. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides generally will have significant amino acid identity relative to each other. A 5 polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

10 THE INVENTION

The invention is based on the discovery of new human transcriptional regulator molecules (HTRM), the polynucleotides encoding HTRM, and the use of these compositions for the diagnosis, treatment, or prevention of cell proliferative and immune disorders.

Table 1 lists the Incyte Clones used to derive full length nucleotide sequences encoding 15 HTRM. Columns 1 and 2 show the sequence identification numbers (SEQ ID NO) of the amino acid and nucleic acid sequences, respectively. Column 3 shows the Clone ID of the Incyte Clone in which nucleic acids encoding each HTRM were identified, and column 4, the cDNA libraries from which these clones were isolated. Column 5 shows Incyte clones, their corresponding cDNA libraries, and shotgun sequences. The clones and shotgun sequences are part of the consensus 20 nucleotide sequence of each HTRM and are useful as fragments in hybridization technologies.

The columns of Table 2 show various properties of the polypeptides of the invention: column 1 references the SEQ ID NO; column 2 shows the number of amino acid residues in each polypeptide; column 3, potential phosphorylation sites; column 4, potential glycosylation sites; column 5, the amino acid residues comprising signature sequences and motifs; column 6, the 25 identity of each protein; and column 7, analytical methods used to identify each protein through sequence homology and protein motifs.

The columns of Table 3 show the tissue-specificity and diseases, disorders, or conditions associated with nucleotide sequences encoding HTRM. The first column of Table 3 lists the nucleotide sequence identifiers. The second column lists tissue categories which express HTRM as 30 a fraction of total tissue categories expressing HTRM. The third column lists the diseases, disorders, or conditions associated with those tissues expressing HTRM. The fourth column lists the vectors used to subclone the cDNA library.

The following fragments of the nucleotide sequences encoding HTRM are useful in hybridization or amplification technologies to identify SEQ ID NO:110-130 and to distinguish 35 between SEQ ID NO:110-130 and related polynucleotide sequences. The useful fragments are the

fragment of SEQ ID NO:110 from about nucleotide 273 to about nucleotide 317; the fragment of SEQ ID NO:111 from about nucleotide 217 to about nucleotide 261 the fragment of SEQ ID NO:112 from about nucleotide 273 to about nucleotide 308; the fragment of SEQ ID NO:113 from about nucleotide 163 to about nucleotide 207; the fragment of SEQ ID NO:114 from about 5 nucleotide 433 to about nucleotide 477; the fragment of SEQ ID NO:115 from about nucleotide 597 to about nucleotide 641; the fragment of SEQ ID NO:116 from about nucleotide 111 to about nucleotide 146; the fragment of SEQ ID NO:117 from about nucleotide 217 to about nucleotide 261; the fragment of SEQ ID NO:118 from about nucleotide 867 to about nucleotide 911; the fragment of SEQ ID NO:119 from about nucleotide 1082 to about nucleotide 1126; the fragment 10 of SEQ ID NO:120 from about nucleotide 702 to about nucleotide 748; the fragment of SEQ ID NO:121 from about nucleotide 380 to about nucleotide 424; the fragment of SEQ ID NO:122 from about nucleotide 352 to about nucleotide 396; the fragment of SEQ ID NO:123 from about nucleotide 219 to about nucleotide 263; the fragment of SEQ ID NO:124 from about nucleotide 326 to about nucleotide 370; the fragment of SEQ ID NO:125 from about nucleotide 595 to about 15 nucleotide 639; the fragment of SEQ ID NO:126 from about nucleotide 272 to about nucleotide 316; the fragment of SEQ ID NO:127 from about nucleotide 163 to about nucleotide 207; the fragment of SEQ ID NO:128 from about nucleotide 271 to about nucleotide 315; the fragment of SEQ ID NO:129 from about nucleotide 866 to about nucleotide 910; and the fragment of SEQ ID NO:130 from about nucleotide 487 to about nucleotide 531.

20 The invention also encompasses HTRM variants. A preferred HTRM variant is one which has at least about 80%, more preferably at least about 90%, and most preferably at least about 95% amino acid sequence identity to the HTRM amino acid sequence, and which contains at least one functional or structural characteristic of HTRM.

The invention also encompasses polynucleotides which encode HTRM. In a particular 25 embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:66-130, which encodes HTRM.

The invention also encompasses a variant of a polynucleotide sequence encoding HTRM. In particular, such a variant polynucleotide sequence will have at least about 70%, more preferably at least about 85%, and most preferably at least about 95% polynucleotide sequence identity to the 30 polynucleotide sequence encoding HTRM. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:66-130 which has at least about 70%, more preferably at least about 85%, and most preferably at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the 35 group consisting of SEQ ID NO:66-130. Any one of the polynucleotide variants described above

can encode an amino acid sequence which contains at least one functional or structural characteristic of HTRM.

- It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding HTRM, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring HTRM, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode HTRM and its variants are preferably capable of hybridizing to the nucleotide sequence of the naturally occurring HTRM under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding HTRM or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding HTRM and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode HTRM and HTRM derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding HTRM or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:66-130 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) Methods Enzymol. 152:399-407; Kimmel, A.R. (1987) Methods Enzymol. 152:507-511.) For example, stringent salt concentration will ordinarily be less than about 750 mM NaCl and 75 mM trisodium citrate, preferably less than about 500 mM NaCl and 50 mM trisodium citrate, and most preferably less than about 250 mM NaCl and 25 mM trisodium citrate. Low stringency hybridization can be obtained in the absence of organic solvent, e.g., formamide, while high stringency hybridization can be obtained in the presence of at least about 35% formamide.

and most preferably at least about 50% formamide. Stringent temperature conditions will ordinarily include temperatures of at least about 30°C, more preferably of at least about 37°C, and most preferably of at least about 42°C. Varying additional parameters, such as hybridization time, the concentration of detergent, e.g., sodium dodecyl sulfate (SDS), and the inclusion or exclusion 5 of carrier DNA, are well known to those skilled in the art. Various levels of stringency are accomplished by combining these various conditions as needed. In a preferred embodiment, hybridization will occur at 30°C in 750 mM NaCl, 75 mM trisodium citrate, and 1% SDS. In a more preferred embodiment, hybridization will occur at 37°C in 500 mM NaCl, 50 mM trisodium citrate, 1% SDS, 35% formamide, and 100 µg/ml denatured salmon sperm DNA (ssDNA). In a 10 most preferred embodiment, hybridization will occur at 42°C in 250 mM NaCl, 25 mM trisodium citrate, 1% SDS, 50 % formamide, and 200 µg/ml ssDNA. Useful variations on these conditions will be readily apparent to those skilled in the art.

The washing steps which follow hybridization can also vary in stringency. Wash 15 stringency conditions can be defined by salt concentration and by temperature. As above, wash stringency can be increased by decreasing salt concentration or by increasing temperature. For example, stringent salt concentration for the wash steps will preferably be less than about 30 mM NaCl and 3 mM trisodium citrate, and most preferably less than about 15 mM NaCl and 1.5 mM trisodium citrate. Stringent temperature conditions for the wash steps will ordinarily include temperature of at least about 25°C, more preferably of at least about 42°C, and most preferably of 20 at least about 68°C. In a preferred embodiment, wash steps will occur at 25°C in 30 mM NaCl, 3 mM trisodium citrate, and 0.1% SDS. In a more preferred embodiment, wash steps will occur at 42°C in 15 mM NaCl, 1.5 mM trisodium citrate, and 0.1% SDS. In a most preferred embodiment, wash steps will occur at 68°C in 15 mM NaCl, 1.5 mM trisodium citrate, and 0.1% SDS.

Additional variations on these conditions will be readily apparent to those skilled in the art.

25 Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASe (US Biochemical, Cleveland OH), Taq polymerase (Perkin-Elmer), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the 30 ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the Hamilton MICROLAB 2200 (Hamilton, Reno NV), Peltier Thermal Cycler 200 (PTC200; MJ Research, Watertown MA) and the ABI CATALYST 800 (Perkin-Elmer). Sequencing is then carried out using either ABI 373 or 377 DNA sequencing systems (Perkin-Elmer) or the MEGABACE 1000 DNA sequencing system 35 (Molecular Dynamics, Sunnyvale CA). The resulting sequences are analyzed using a variety of

algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853.)

- The nucleic acid sequences encoding HTRM may be extended utilizing a partial
- 5 nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) PCR Methods Appl. 2:318-322.) Another method, inverse PCR, uses primers that extend in divergent
 - 10 directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) PCR Methods Appl. 1:111-119.) In
 - 15 this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) Nucleic Acids Res. 19:3055-306). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic
 - 20 DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 Primer Analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.
 - 25 When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. In addition, random-primed libraries, which often include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.
 - 30 Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal
 - 35 using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, Perkin-Elmer),

and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof
5 which encode HTRM may be cloned in recombinant DNA molecules that direct expression of HTRM, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and used to express HTRM.

10 The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter HTRM-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example,
15 oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

In another embodiment, sequences encoding HTRM may be synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) Nucl. Acids Res. Symp. Ser. 215-223, and Horn, T. et al. (1980) Nucl. Acids Res. Symp. Ser. 225-232.) Alternatively, HTRM itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solid-phase techniques. (See, e.g., Roberge, J.Y. et al. (1995) Science 269:202-204.) Automated synthesis may be achieved using the ABI 431A Peptide Synthesizer (Perkin-Elmer). Additionally, the amino acid sequence of
25 HTRM, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography. (See, e.g., Chieze, R.M. and F.Z. Regnier (1990) Methods Enzymol. 182:392-421.) The composition of the synthetic peptides may be confirmed by amino acid analysis or by
30 sequencing. (See, e.g., Creighton, T. (1984) Proteins, Structures and Molecular Properties. WH Freeman, New York N.Y.)

In order to express a biologically active HTRM, the nucleotide sequences encoding HTRM or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted
35 coding sequence in a suitable host. These elements include regulatory sequences, such as

enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences encoding HTRM. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding HTRM. Such signals include the ATG initiation codon and adjacent

- 5 sequences, e.g. the Kozak sequence. In cases where sequences encoding HTRM and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous
10 translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used. (See, e.g., Scharf, D. et al. (1994) *Results Probl. Cell Differ.* 20:125-162.)

Methods which are well known to those skilled in the art may be used to construct
15 expression vectors containing sequences encoding HTRM and appropriate transcriptional and translational control elements. These methods include *in vitro* recombinant DNA techniques, synthetic techniques, and *in vivo* genetic recombination. (See, e.g., Sambrook, J. et al. (1989) *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) *Current Protocols in Molecular Biology*, John Wiley & Sons, 20 New York NY, ch. 9, 13, and 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences encoding HTRM. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral
25 expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected
30 depending upon the use intended for polynucleotide sequences encoding HTRM. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding HTRM can be achieved using a multifunctional *E. coli* vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or pSPORT1 plasmid (Life Technologies). Ligation of sequences encoding HTRM into the vector's multiple cloning site disrupts the lacZ gene, allowing a colorimetric screening procedure
35 for identification of transformed bacteria containing recombinant molecules. In addition, these

vectors may be useful for *in vitro* transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heege, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509.) When large quantities of HTRM are needed, e.g. for the production of antibodies, vectors which direct high level expression of HTRM 5 may be used. For example, vectors containing the strong, inducible T5 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of HTRM. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH, may be used in the yeast Saccharomyces cerevisiae or Pichia pastoris. In addition, such vectors 10 direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, supra; Grant et al. (1987) Methods Enzymol. 153:516-54; and Scorer, C. A. et al. (1994) Bio/Technology 12:181-184.)

Plant systems may also be used for expression of HTRM. Transcription of sequences 15 encoding HTRM may be driven viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) EMBO J. 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, G. et al. (1984) EMBO J. 3:1671-1680; Broglie, R. et al. (1984) Science 224:838-843; and Winter, J. et al. (1991) Results Probl. Cell 20 Differ. 17:85-105.) These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding HTRM may be ligated 25 into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses HTRM in host cells. (See, e.g., Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. 30 SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments 35 of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) Nat Genet. 15:345-355.)

For long term production of recombinant proteins in mammalian systems, stable expression of HTRM in cell lines is preferred. For example, sequences encoding HTRM can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

- 10 Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk*' or *apr* cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dfr* confers 15 resistance to methotrexate; *neo* confers resistance to the aminoglycosides, neomycin and G-418; and *als* or *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., 20 Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech), β glucuronidase and its substrate β -glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) Methods Mol. 25 Biol. 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding HTRM is inserted within a marker gene sequence, transformed cells containing sequences encoding HTRM can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding HTRM under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the nucleic acid sequence encoding HTRM and that express HTRM may be identified by a variety of procedures known to those of skill in the art. 35 These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR

amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

- Immunological methods for detecting and measuring the expression of HTRM using either
- 5 specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on HTRM is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art.
- 10 (See, e.g., Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St Paul MN, Sect. IV; Coligan, J. E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ).

- A wide variety of labels and conjugation techniques are known by those skilled in the art
- 15 and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding HTRM include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding HTRM, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are
- 20 commercially available, and may be used to synthesize RNA probes *in vitro* by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by Amersham Pharmacia Biotech, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for ease of detection include radionuclides, enzymes,
- 25 fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

- Host cells transformed with nucleotide sequences encoding HTRM may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the
- 30 sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode HTRM may be designed to contain signal sequences which direct secretion of HTRM through a prokaryotic or eukaryotic cell membrane.

- In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications
- 35 of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation.

phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38), are available from 5 the American Type Culture Collection (ATCC, Bethesda MD) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding HTRM may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric HTRM protein 10 containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of HTRM activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, 15 FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be 20 engineered to contain a proteolytic cleavage site located between the HTRM encoding sequence and the heterologous protein sequence, so that HTRM may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel (1995, *supra*, ch 10). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

25 In a further embodiment of the invention, synthesis of radiolabeled HTRM may be achieved *in vitro* using the TNT rabbit reticulocyte lysate or wheat germ extract systems (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, preferably ^{35}S -methionine.

30 Fragments of HTRM may be produced not only by recombinant production, but also by direct peptide synthesis using solid-phase techniques. (See, e.g., Creighton, *supra*, pp. 55-60.) Protein synthesis may be performed by manual techniques or by automation. Automated synthesis may be achieved, for example, using the ABI 431A Peptide Synthesizer (Perkin-Elmer). Various fragments of HTRM may be synthesized separately and then combined to produce the full length 35 molecule.

THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of HTRM and human transcriptional regulator molecules. In addition, the expression of HTRM is closely associated with cell proliferation, inflammation, and the immune response. Therefore, HTRM appears to play a role in cell proliferative and immune disorders. In the treatment of disorders associated with increased HTRM expression or activity, it is desirable to decrease the expression or activity of HTRM. In the treatment of disorders associated with decreased HTRM expression or activity, it is desirable to increase the expression or activity of HTRM.

- Therefore, in one embodiment, HTRM or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of HTRM. Examples of such disorders include, but are not limited to, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia; cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; and an immune disorder such as acquired immunodeficiency syndrome (AIDS). Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma

In another embodiment, a vector capable of expressing HTRM or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of HTRM including, but not limited to, those described above.

In a further embodiment, a pharmaceutical composition comprising a substantially purified HTRM in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of HTRM including, but not limited to, those provided above.

5 In still another embodiment, an agonist which modulates the activity of HTRM may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of HTRM including, but not limited to, those listed above.

In a further embodiment, an antagonist of HTRM may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of HTRM. Examples of

10 such disorders include, but are not limited to, those described above. In one aspect, an antibody which specifically binds HTRM may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissue which express HTRM.

In an additional embodiment, a vector expressing the complement of the polynucleotide
15 encoding HTRM may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of HTRM including, but not limited to, those described above.

In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary sequences, or vectors of the invention may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination
20 therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

25 An antagonist of HTRM may be produced using methods which are generally known in the art. In particular, purified HTRM may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind HTRM. Antibodies to HTRM may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments,
30 and fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are especially preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with HTRM or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various
35 adjuvants may be used to increase immunological response. Such adjuvants include, but are not

limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol.

Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

- 5 It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to HTRM have an amino acid sequence consisting of at least about 5 amino acids, and, more preferably, of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein and contain the entire amino acid sequence of a small, naturally occurring molecule. Short stretches of
10 HTRM amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to HTRM may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-
15 hybridoma technique. (See, e.g., Kohler, G. et al. (1975) *Nature* 256:495-497; Kozbor, D. et al. (1985) *J. Immunol. Methods* 81:31-42; Cote, R.J. et al. (1983) *Proc. Natl. Acad. Sci.* 80:2026-2030; and Cole, S.P. et al. (1984) *Mol. Cell Biol.* 62:109-120.)

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate
20 antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) *Proc. Natl. Acad. Sci.* 81:6851-6855; Neuberger, M.S. et al. (1984) *Nature* 312:604-608; and Takeda, S. et al. (1985) *Nature* 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce HTRM-specific single chain antibodies. Antibodies with related specificity, but of distinct
25 idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton D.R. (1991) *Proc. Natl. Acad. Sci.* 88:10134-10137.)

Antibodies may also be produced by inducing *in vivo* production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) *Proc. Natl. Acad. Sci.* 86:
30 3833-3837; Winter, G. et al. (1991) *Nature* 349:293-299.)

Antibody fragments which contain specific binding sites for HTRM may also be generated. For example, such fragments include, but are not limited to, F(ab')2 fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the F(ab')2 fragments. Alternatively, Fab expression libraries may be
35 constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired

specificity. (See, e.g., Huse, W.D. et al. (1989) *Science* 246:1275-1281.)

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between HTRM and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering HTRM epitopes is preferred, but a competitive binding assay may also be employed (Pound, *supra*).

- Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for HTRM. Affinity is expressed as an association constant, K_a , which is defined as the molar concentration of HTRM-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The K_a determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple HTRM epitopes, represents the average affinity, or avidity, of the antibodies for HTRM. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular HTRM epitope, represents a true measure of affinity. High-affinity antibody preparations with K_a ranging from about 10^9 to 10^{12} L/mole are preferred for use in immunoassays in which the HTRM-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with K_a ranging from about 10^6 to 10^7 L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of HTRM, preferably in active form, from the antibody (Catty, D. (1988) *Antibodies, Volume I: A Practical Approach*. IRL Press, Washington, DC; Liddell, J. E. and Cryer, A. (1991) *A Practical Guide to Monoclonal Antibodies*, John Wiley & Sons, New York NY).

- The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is preferred for use in procedures requiring precipitation of HTRM-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available.
- (See, e.g., Catty, *supra*, and Coligan et al. *supra*.)

In another embodiment of the invention, the polynucleotides encoding HTRM, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, the complement of the polynucleotide encoding HTRM may be used in situations in which it would be desirable to block the transcription of the mRNA. In particular, cells may be transformed with sequences complementary to polynucleotides encoding HTRM. Thus, complementary molecules

or fragments may be used to modulate HTRM activity, or to achieve regulation of gene function. Such technology is now well known in the art, and sense or antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding HTRM.

5 Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. Methods which are well known to those skilled in the art can be used to construct vectors to express nucleic acid sequences complementary to the polynucleotides encoding HTRM. (See, e.g., Sambrook, *supra*; Ausubel, 1995, *supra*.)

10 Genes encoding HTRM can be turned off by transforming a cell or tissue with expression vectors which express high levels of a polynucleotide, or fragment thereof, encoding HTRM. Such constructs may be used to introduce untranslatable sense or antisense sequences into a cell. Even in the absence of integration into the DNA, such vectors may continue to transcribe RNA molecules until they are disabled by endogenous nucleases. Transient expression may last for a 15 month or more with a non-replicating vector, and may last even longer if appropriate replication elements are part of the vector system.

As mentioned above, modifications of gene expression can be obtained by designing complementary sequences or antisense molecules (DNA, RNA, or PNA) to the control, 5', or regulatory regions of the gene encoding HTRM. Oligonucleotides derived from the transcription 20 initiation site, e.g., between about positions -10 and +10 from the start site, are preferred. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature. (See, e.g., Gee, J.E. et al. 25 (1994) in Huber, B.E. and B.I. Carr, *Molecular and Immunologic Approaches*, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage 30 of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding HTRM.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: 35 GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20

ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

- 5 Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding HTRM. Such DNA sequences may be
10 incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

- RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3'
15 ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms
20 of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

- Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers
25 may be achieved using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) *Nature Biotechnology* 15:462-466.)

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as dogs, cats, cows, horses, rabbits, monkeys, and most preferably, humans.

- 30 An additional embodiment of the invention relates to the administration of a pharmaceutical or sterile composition, in conjunction with a pharmaceutically acceptable carrier, for any of the therapeutic effects discussed above. Such pharmaceutical compositions may consist of HTRM, antibodies to HTRM, and mimetics, agonists, antagonists, or inhibitors of HTRM. The compositions may be administered alone or in combination with at least one other agent, such as a
35 stabilizing compound, which may be administered in any sterile, biocompatible pharmaceutical

carrier including, but not limited to, saline, buffered saline, dextrose, and water. The compositions may be administered to a patient alone, or in combination with other agents, drugs, or hormones.

The pharmaceutical compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial,

- 5 intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used

- 10 pharmaceutically. Further details on techniques for formulation and administration may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA).

Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as tablets, 15 pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

Pharmaceutical preparations for oral use can be obtained through combining active compounds with solid excipient and processing the resultant mixture of granules (optionally, after grinding) to obtain tablets or dragee cores. Suitable auxiliaries can be added, if desired. Suitable 20 excipients include carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, and sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxypropylmethyl-cellulose, or sodium carboxymethylcellulose; gums, including arabic and tragacanth; and proteins, such as gelatin and collagen. If desired, disintegrating or solubilizing agents may be added, such as the cross-linked polyvinyl pyrrolidone, 25 agar, and alginic acid or a salt thereof, such as sodium alginate.

Dragee cores may be used in conjunction with suitable coatings, such as concentrated sugar solutions, which may also contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for 30 product identification or to characterize the quantity of active compound, i.e., dosage.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol. Push-fit capsules can contain active ingredients mixed with fillers or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft 35 capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty

oils, liquid, or liquid polyethylene glycol with or without stabilizers.

- Pharmaceutical formulations suitable for parenteral administration may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils, such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate, triglycerides, or liposomes. Non-lipid polycationic amino polymers may also be used for delivery. Optionally, the suspension may also contain suitable stabilizers or agents to increase the solubility of the compounds and allow for the preparation of highly concentrated solutions.

For topical or nasal administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

- The pharmaceutical compositions of the present invention may be manufactured in a manner that is known in the art, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes.

The pharmaceutical composition may be provided as a salt and can be formed with many acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, and succinic acid. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms. In other cases, the preferred preparation may be a lyophilized powder which may contain any or all of the following: 1 mM to 50 mM histidine, 0.1% to 2% sucrose, and 2% to 7% mannitol, at a pH range of 4.5 to 5.5, that is combined with buffer prior to use.

- After pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. For administration of HTRM, such labeling would include amount, frequency, and method of administration.

Pharmaceutical compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

- For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells or in animal models such as mice, rats, rabbits, dogs, or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example HTRM or fragments thereof, antibodies of HTRM, and agonists, antagonists or inhibitors of HTRM, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals,

5 such as by calculating the ED₅₀ (the dose therapeutically effective in 50% of the population) or LD₅₀ (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, and it can be expressed as the LD₅₀/ED₅₀ ratio. Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The

10 dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED₅₀ with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of

15 the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance

20 rate of the particular formulation.

Normal dosage amounts may vary from about 0.1 µg to 100,000 µg, up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their

25 inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

DIAGNOSTICS

In another embodiment, antibodies which specifically bind HTRM may be used for the diagnosis of disorders characterized by expression of HTRM, or in assays to monitor patients

30 being treated with HTRM or agonists, antagonists, or inhibitors of HTRM. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for HTRM include methods which utilize the antibody and a label to detect HTRM in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter

35 molecule. A wide variety of reporter molecules, several of which are described above, are known

in the art and may be used.

A variety of protocols for measuring HTRM, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of HTRM expression. Normal or standard values for HTRM expression are established by combining body

- 5 fluids or cell extracts taken from normal mammalian subjects, preferably human, with antibody to HTRM under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, preferably by photometric means. Quantities of HTRM expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for
10 diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding HTRM may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantitate gene expression in biopsied tissues in which expression of HTRM
15 may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of HTRM, and to monitor regulation of HTRM levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding HTRM or closely related
20 molecules may be used to identify nucleic acid sequences which encode HTRM. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification (maximal, high, intermediate, or low), will determine whether the probe identifies only naturally occurring sequences encoding HTRM, allelic variants, or related sequences.

25 Probes may also be used for the detection of related sequences, and should preferably have at least 50% sequence identity to any of the HTRM encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:66-130 or from genomic sequences including promoters, enhancers, and introns of the HTRM gene.

30 Means for producing specific hybridization probes for DNAs encoding HTRM include the cloning of polynucleotide sequences encoding HTRM or HTRM derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a
35 variety of reporter groups, for example, by radionuclides such as ³²P or ³⁵S, or by enzymatic labels,

- such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.
- Polynucleotide sequences encoding HTRM may be used for the diagnosis of disorders associated with expression of HTRM. Examples of such disorders include, but are not limited to, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis,
- 5 cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia; cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas,
- 10 parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; and an immune disorder such as acquired immunodeficiency syndrome (AIDS). Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes
- 15 mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma,
- 20 Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma. The polynucleotide sequences encoding HTRM may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR
- 25 technologies; in dipstick, pin, and multifORMAT ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered HTRM expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding HTRM may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The

30 nucleotide sequences encoding HTRM may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantitated and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide

35 sequences encoding HTRM in the sample indicates the presence of the associated disorder. Such

assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

- In order to provide a basis for the diagnosis of a disorder associated with expression of HTRM, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding HTRM, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

- Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

- With respect to cancer, the presence of an abnormal amount of transcript (either under- or overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

- Additional diagnostic uses for oligonucleotides designed from the sequences encoding HTRM may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced *in vitro*. Oligomers will preferably contain a fragment of a polynucleotide encoding HTRM, or a fragment of a polynucleotide complementary to the polynucleotide encoding HTRM, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantitation of closely related DNA or RNA sequences.

- Methods which may also be used to quantitate the expression of HTRM include radiolabeling or biotinylation nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. (See, e.g., Melby, P.C. et al. (1993) J. Immunol. Methods 159:235-244; Duplaa, C. et al. (1993) Anal. Biochem. 229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in an ELISA format where the oligomer of interest is presented in various dilutions and a spectrophotometric or

colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as targets in a microarray. The microarray can be used to monitor the expression level of large numbers of genes simultaneously and to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, and to develop and monitor the activities of therapeutic agents.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci. 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.)

In another embodiment of the invention, nucleic acid sequences encoding HTRM may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) Nat Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127-134; and Trask, B.J. (1991) Trends Genet. 7:149-154.)

Fluorescent *in situ* hybridization (FISH) may be correlated with other physical chromosome mapping techniques and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, *supra*, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) site. Correlation between the location of the gene encoding HTRM on a physical chromosomal map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder. The nucleotide sequences of the invention may be used to detect differences in gene sequences among normal, carrier, and affected individuals.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the number or arm of a particular human chromosome is not known. New sequences can be assigned to chromosomal arms by physical mapping. This provides valuable information to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the disease or syndrome has been

crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) *Nature* 336:577-580.) The nucleotide sequence of the subject invention may also be used to detect differences in the chromosomal
5 location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

In another embodiment of the invention, HTRM, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of
10 binding complexes between HTRM and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with HTRM, or
15 fragments thereof, and washed. Bound HTRM is then detected by methods well known in the art. Purified HTRM can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which
20 neutralizing antibodies capable of binding HTRM specifically compete with a test compound for binding HTRM. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with HTRM.

In additional embodiments, the nucleotide sequences which encode HTRM may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely
25 on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following preferred specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of
30 the remainder of the disclosure in any way whatsoever.

The entire disclosure of all applications, patents, and publications, cited above and below, and of US provisional applications 60/084,254 (filed May 5, 1998), 60/095,827 (filed August 7, 1998), and 60/102,745 (filed Oct. 2, 1998) are hereby incorporated by reference.

EXAMPLES

35 I. Construction of cDNA Libraries

RNA was purchased from Clontech or isolated from tissues described in Table 4. Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Life Technologies), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In some cases, RNA was treated with DNase. For most libraries, poly(A+) RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN, Valencia CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERSCRIPT plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, *supra*, units 5.1-6.6). Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., pBLUESCRIPT plasmid (Stratagene), pSPORT1 plasmid (Life Technologies), or pINCY (Incyte Pharmaceuticals, Palo Alto CA). Recombinant plasmids were transformed into competent *E. coli* cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5 α , DH10B, or ElectroMAX DH10B from Life Technologies.

II. Isolation of cDNA Clones

Plasmids were recovered from host cells by *in vivo* excision, using the UNIZAP vector system (Stratagene) or cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the REAL Prep 96 plasmid kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) *Anal. Biochem.* 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified 5 fluorometrically using PICOGREEN dye (Molecular Probes, Eugene OR) and a Fluoroskan II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

III. Sequencing and Analysis

The cDNAs were prepared for sequencing using the ABI CATALYST 800 (Perkin-Elmer) or the HYDRA microdispenser (Robbins Scientific) or MICROLAB 2200 (Hamilton) systems in 10 combination with the PTC-200 thermal cyclers (MJ Research). The cDNAs were sequenced using the ABI PRISM 373 or 377 sequencing systems (Perkin-Elmer) and standard ABI protocols, base calling software, and kits. In one alternative, cDNAs were sequenced using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics). In another alternative, the cDNAs were 15 amplified and sequenced using the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Perkin-Elmer). In yet another alternative, cDNAs were sequenced using solutions and dyes from Amersham Pharmacia Biotech. Reading frames for the ESTs were determined using standard methods (reviewed in Ausubel, 1997, *supra*, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example V.

The polynucleotide sequences derived from cDNA, extension, and shotgun sequencing 20 were assembled and analyzed using a combination of software programs which utilize algorithms well known to those skilled in the art. Table 5 summarizes the software programs, descriptions, references, and threshold parameters used. The first column of Table 5 shows the tools, programs, and algorithms used, the second column provides a brief description thereof, the third column presents the references which are incorporated by reference herein, and the fourth column 25 presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the probability the greater the homology). Sequences were analyzed using MACDNASIS PRO software (Hitachi Software Engineering, S. San Francisco CA) and LASERGENE software (DNASTAR).

cDNAs were also compared to sequences in GenBank using a search algorithm developed 30 by Applied Biosystems and incorporated into the INHERIT™ 670 sequence analysis system. In this algorithm, Pattern Specification Language (TRW Inc, Los Angeles, CA) was used to determine regions of homology. The three parameters that determine how the sequence comparisons run were window size, window offset, and error tolerance. Using a combination of these three parameters, the DNA database was searched for sequences containing regions of 35 homology to the query sequence, and the appropriate sequences were scored with an initial value.

Subsequently, these homologous regions were examined using dot matrix homology plots to distinguish regions of homology from chance matches. Smith-Waterman alignments were used to display the results of the homology search.

Peptide and protein sequence homologies were ascertained using the INHERIT- 670
5 sequence analysis system using the methods similar to those used in DNA sequence homologies. Pattern Specification Language and parameter windows were used to search protein databases for sequences containing regions of homology which were scored with an initial value. Dot-matrix homology plots were examined to distinguish regions of significant homology from chance matches.

10 The polynucleotide sequences were validated by removing vector, linker, and polyA sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programing, and dinucleotide nearest neighbor analysis. The sequences were then queried against a selection of public databases such as GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS to acquire annotation, using programs based on
15 BLAST, FASTA, and BLIMPS. The sequences were assembled into full length polynucleotide sequences using programs based on Phred, Phrap, and Consed, and were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length amino acid sequences, and these full length sequences were subsequently analyzed by querying against
20 databases such as the GenBank databases (described above), SwissProt, BLOCKS, PRINTS, PFAM, and Prosite.

The programs described above for the assembly and analysis of full length polynucleotide and amino acid sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:110-130. Fragments from about 20 to about 4000 nucleotides which are useful in
25 hybridization and amplification technologies were described in The Invention section above.

IV. Northern Analysis

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound. (See, e.g., Sambrook, supra, ch. 7; 30 Ausubel, 1995, supra, ch. 4 and 16.)

Analogous computer techniques applying BLAST were used to search for identical or related molecules in nucleotide databases such as GenBank or LIFESEQ database (Incyte Pharmaceuticals). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any
35 particular match is categorized as exact or similar. The basis of the search is the product score,

which is defined as:

$$\frac{\% \text{ sequence identity} \times \% \text{ maximum BLAST score}}{100}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. For example, with a product score of 40, the match will be exact within a 1% to 2% error, and, with a product score of 70, the match will be exact. Similar molecules are usually identified by selecting those which show product scores between 15 and 40, although lower scores may identify related molecules.

The results of northern analyses are reported a percentage distribution of libraries in which the transcript encoding HTRM occurred. Analysis involved the categorization of cDNA libraries by organ/tissue and disease. The organ/tissue categories included cardiovascular, dermatologic, developmental, endocrine, gastrointestinal, hematopoietic/immune, musculoskeletal, nervous, reproductive, and urologic. The disease categories included cancer, inflammation/trauma, fetal, neurological, and pooled. For each category, the number of libraries expressing the sequence of interest was counted and divided by the total number of libraries across all categories. Percentage values of tissue-specific and disease expression are reported in Table 3.

V. Extension of HTRM Encoding Polynucleotides

The full length nucleic acid sequence of SEQ ID NO:66-130 was produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer, to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing Mg²⁺, (NH₄)₂SO₄, and β-mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+

were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100 µl PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5 µl of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the concentration of DNA. A 5 µl to 10 µl aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose mini-gel to determine which reactions were 10 successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) 15 agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent *E. coli* cells. Transformed cells were selected on antibiotic-containing media, individual colonies were picked and cultured overnight at 20 37°C in 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was 25 quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethylsulphoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Perkin-Elmer).

30 In like manner, the nucleotide sequence of SEQ ID NO:66-130 is used to obtain 5' regulatory sequences using the procedure above, oligonucleotides designed for such extension, and an appropriate genomic library.

VI. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:66-130 are employed to screen cDNAs, 35 genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20

base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 μ Ci of [γ -³²P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase

- 5 (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10^7 counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Asc I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

10 The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under increasingly stringent conditions up to 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. After XOMAT-AR film (Eastman Kodak, Rochester NY) is exposed to the blots to film
15 for several hours, hybridization patterns are compared visually.

VII. Microarrays

A chemical coupling procedure and an ink jet device can be used to synthesize array elements on the surface of a substrate. (See, e.g., Baldeschweiler, *supra*.) An array analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using
20 thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced by hand or using available methods and machines and contain any appropriate number of elements. After hybridization, nonhybridized probes are removed and a scanner used to determine the levels and patterns of fluorescence. The degree of complementarity and the relative abundance of each probe which hybridizes to an element on the microarray may be assessed through analysis of the
25 scanned images.

Full-length cDNAs, Expressed Sequence Tags (ESTs), or fragments thereof may comprise the elements of the microarray. Fragments suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). Full-length cDNAs, ESTs, or fragments thereof corresponding to one of the nucleotide sequences of the
30 present invention, or selected at random from a cDNA library relevant to the present invention, are arranged on an appropriate substrate, e.g., a glass slide. The cDNA is fixed to the slide using, e.g., UV cross-linking followed by thermal and chemical treatments and subsequent drying. (See, e.g., Schena, M. et al. (1995) Science 270:467-470; Shalon, D. et al. (1996) Genome Res. 6:639-645.) Fluorescent probes are prepared and used for hybridization to the elements on the substrate. The
35 substrate is analyzed by procedures described above.

VIII. Complementary Polynucleotides

Sequences complementary to the HTRM-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring HTRM. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of HTRM. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the 10 HTRM-encoding transcript.

IX. Expression of HTRM

Expression and purification of HTRM is achieved using bacterial or virus-based expression systems. For expression of HTRM in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *npp-lac* (*tac*) hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express HTRM upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of HTRM in eukaryotic cells is achieved by 20 infecting insect or mammalian cell lines with recombinant *Autographica californica* nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding HTRM by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription.

25 Recombinant baculovirus is used to infect *Spodoptera frugiperda* (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E. K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

In most expression systems, HTRM is synthesized as a fusion protein with, e.g.,

- 30 glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from *Schistosoma japonicum*, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be 35 proteolytically cleaved from HTRM at specifically engineered sites. FLAG, an 8-amino acid

peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, *supra*, ch 10 and 16). Purified HTRM obtained by these methods can be used directly in the following activity assay.

X. Demonstration of HTRM Activity

HTRM activity is measured by its ability to stimulate transcription of a reporter gene, essentially as described in Liu, H.Y., et al (1997; EMBO J. 16:5289-5298.). The assay entails the use of a well characterized reporter gene construct, LexA_{op}-LacZ, that consists of LexA DNA transcriptional control elements (LexA_{op}) fused to sequences encoding the *E. coli* β-galactosidase enzyme (LacZ). The methods for fusion gene construction, expression, and introduction into cells, and measurement of β-galactosidase enzyme activity, are well known to those skilled in the art. Sequences encoding HTRM are cloned into a plasmid that directs the synthesis of a fusion protein, LexA-HTRM, consisting of HTRM and a DNA binding domain derived from the LexA transcription factor. The plasmid encoding the LexA-HTRM fusion protein is introduced into yeast cells along with the plasmid containing the LexA_{op}-LacZ reporter gene. The amount of β-galactosidase enzyme activity associated with LexA-HTRM transfected cells, relative to control cells, is proportional to the amount of transcription stimulated by the HTRM gene product.

20 XI. Functional Assays

HTRM function is assessed by expressing the sequences encoding HTRM at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include pCMV SPORT (Life Technologies) and pCR3.1 (Invitrogen, Carlsbad CA), both of which contain the cytomegalovirus promoter. 5-10 µg of recombinant vector are transiently transfected into a human cell line, preferably of endothelial or hematopoietic origin, using either liposome formulations or electroporation. 1-2 µg of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP, and to evaluate properties, for example, their apoptotic state. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA

content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M. G. (1994) *Flow Cytometry*, Oxford, New York NY.

- The influence of HTRM on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding HTRM and either CD64 or CD64-GFP.
- 10 CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding HTRM and other genes of interest can 15 be analyzed by northern analysis or microarray techniques.

XII. Production of HTRM Specific Antibodies

HTRM substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) *Methods Enzymol.* 182:488-495), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols.

- 20 Alternatively, the HTRM amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel, 1995, *supra*, ch. II.)
- 25 Typically, oligopeptides 15 residues in length are synthesized using an ABI 431A Peptide Synthesizer (Perkin-Elmer) using fmoc-chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, *supra*.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for 30 antipeptide activity by, for example, binding the peptide to plastic, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

XIII. Purification of Naturally Occurring HTRM Using Specific Antibodies

- Naturally occurring or recombinant HTRM is substantially purified by immunoaffinity chromatography using antibodies specific for HTRM. An immunoaffinity column is constructed 35 by covalently coupling anti-HTRM antibody to an activated chromatographic resin, such as

CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing HTRM are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of HTRM (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/HTRM binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and HTRM is collected.

XIV. Identification of Molecules Which Interact with HTRM

- HTRM, or biologically active fragments thereof, are labeled with ^{125}I Bolton-Hunter reagent. (See, e.g., Bolton et al. (1973) Biochem. J. 133:529.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled HTRM, washed, and any wells with labeled HTRM complex are assayed. Data obtained using different concentrations of HTRM are used to calculate values for the number, affinity, and association of HTRM with the candidate molecules.
- Various modifications and variations of the described methods and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

Table 1

| Protein SEQ ID NO: | Nucleotide SEQ ID NO: | Clone ID | Library | Fragments |
|-----------------------|--------------------------|----------|-----------|---|
| 1 | 66 | 001106 | U937R0701 | 001106 (U937R0701), 1291142 (BRAN0711), 2590425 (LUNG0722), 1300570 (BRSTND07) |
| 2 | 67 | 004566 | HMC1N0701 | 004566 (HMC1N0701), 3899843 (BRSTUT16), 1412988 (BEPIND01), 788995 (PFSUT03), 505475 (LUNGNC15) |
| 3 | 67 | 052927 | FIBRN0701 | 052927 (FIBRN0701), 2518848 (BRAUT021), 3520218 (LUNG0703), 086878 (LIVRNC0701) |
| 4 | 69 | 082843 | HUVESTB01 | 082843 (HUVESTB01), 4008105 (ENDCNT04), 2083528 (UTSNS0708), 2345764 (HESTUT02), 3771780 (BRSTNC0705), 190782 (CONUT01), 2202559 (SPINNEY02), 2509193 (CONUT01) |
| 5 | 70 | 322349 | EVSIIET02 | 322249 (BOSIET02), 3686018 (HEBAND01), 1853592 (LUNGET03), 815966 (CVART01), 1505002 (BRAUT0707), 1511883 (LUNGNC14), 2232826 (PFSNC016) |
| 6 | 71 | 397653 | PITUN0702 | 397653 (PITUN0702), 491141 (HNT2AGT01), 3809879 (CONUT01), 3562349 (SKN0705), 11518413 (BLADUT06), 3600151 (DRGNC0701), 2474103 (THBNC0703), 2105304 (BRAUT0703), 2187330 (PROS0726), 1781572 (TFANN0702), 2056358 (BEPIND0701), 1888065 (BLADUT07) |
| 7 | 72 | 673766 | CRBLN0701 | 673766 (CRBLN0701), 2494421 (ADRET0701), 3267748 (BRAN0701), 2194042 (THFR0703), 3166455 (RHYN0704), 1712336 (PROSON016), 1844092 (CLUN0704), 1602283 (BLADNC0703), 0313357 (THP1PROB01), 1995828 (BISUT0703), 1495594 (ORPN0702) |
| 8 | 73 | 1504753 | BRAUT07 | 1514753 (BRAUT07), 633939 (NEBUT01), 998401 (KIDN0701), 2959561 (ATREN0709), 3483904 (KIDN0731), 2741379 (BRSTUT14), 1965504 (BISUT0704), 5865535 (UTSNS0702) |
| 9 | 74 | 1760185 | PITUN0703 | 1760085 (PITUN0703), 1914471 (PROSUT04), 836811 (PROSN0701), 729798 (LUNGNC0703), 1250847 (BRAIN0711), 1492387 (PROSKN01), 1366472 (SOPNC0702) |

Table 1 cont.

| Protein SEQ ID NO: | Nucleotide SEQ ID NO: | Clone ID | Library | Fragments |
|-----------------------|--------------------------|----------|-----------|--|
| 10 | 75 | 1805061 | SINTNOT13 | 1805061 (SINTNOT13), 1435949 (PANCNOT09), 086122 (LIVNOT01) 1482366 (CORENOT02), 1835310 (BRAINNOT01), 1333758 (COLNOT13), 3521449 (LUNGNOT03) |
| 11 | 76 | 1850120 | LUNGNOT03 | 1850120 (LUNGNOT03), 3126350 (LUNGNOT12), 786916 (PROSNOT05) 2899740 (BRCNOT01), 1259221 (HEMNOT03), 1334740 (COLNOT13), 2466350 (HYRNOT08) |
| 12 | 77 | 1852230 | LUNGNOT03 | 1852230 (LUNGNOT03), 2901081 (DGCNOT01), 1384842 (BRAITUT08) 129354 (PRAMNOT03), 1964126 (BISTNOT04) |
| 13 | 78 | 1944530 | PITNOT01 | 1944530 (PTENNOT01), 2608142 and 2809196 (BLADUT08), 2961779 (ADENOT09) |
| 14 | 79 | 2019742 | CORNOT01 | 2019742 (CORNOT01), 2368014 (SCORN04), 166472 (LIVNOT01) 1875993 (LEURNOT02), 1522480 (BLADUT04), 1418495 (KIDRNOT09), 149730 (FEBRNOT02) |
| 15 | 80 | 2056042 | BEPNOT01 | 2056042 (BEPNOT01), 3097391 (CERVNOT03), 1985203 (LUNGAST01), 1962319 (BISTNOT04), 135716 (COLNOT13) |
| 16 | 81 | 2398682 | THPIAZT01 | 2398682 (THPIAZT01), 159706 (ADERIN01), 2443910 (THPINOT03) 2382189 (ISLNOT01), 2268661 (BRAINNOT01), 1864422 (PROSNOT19) |
| 17 | 82 | 2518753 | BRAITUT21 | 2518753 (BRAITUT21), 4001219 (HNTZAS07), 2606361 (LUNGNOT07), 449043 (FLYNOT02), SAE01190 |
| 18 | 83 | 2709055 | PONSAZT01 | 2709055 (PONSAZT01), 2309703 (NGANNOT01), 1661042 (UREPNTU01) 27161284 (EGSNOT02), 2495934 (THPIAZT01), SBLA01183, SBLA00549, SBLA00975 |
| 19 | 84 | 2724537 | LUNGUT10 | 2724537 (LUNGUT10) 3865823 (BMRNTO3), 952779 (SCORN001), 2049127 (LIVNOT02), 1824284 (BLATNOT01), 1870588 and 1869666 (SKINNOT01), 2526505 (PROSNOT12), SAE01744 SAE01045, SAPRO04072, SAPRO0149 |

Table 1 cont.

| Protein SEQ ID NO: | Nucleotide SEQ ID NO: | Clone ID | Library | Fragment |
|-----------------------|--------------------------|----------|------------|--|
| 20 | 85 | 025818 | SPUNFETO1 | 025818H1, 025818X12, and 025818X3 (SPUNFETO1), 783259H1 (MYOMRNOTO01), 82616SR1 (PRONNOT06) |
| 21 | 86 | 438283 | THYRNOT01 | 43828H1 and 438283X29 (THYRNOT01), SAGA01136F1, SAGA01136R1, SAGA01136P1, SAGA01671F1, SAGA01671R1, SAGA01671P1, S22Z01038R1 |
| 22 | 87 | 619659 | PGANNOT01 | 619694H1, 619694X11, and 619699X18 (PGANNOT01), 646194T6 (BFSUTU02), 134230520F1 (BLADNOT04), 1724316F6 (PROSNOT14) |
| 23 | 88 | 693452 | SYNORAT03 | 11814H1 (MUSCNNOT01), 693452H1 and 693452R6 (SYNORAT03), 2455518F6 and 2455518H1 (ENDANOT01), 4500313H1 (BRAVXTT02) |
| 24 | 89 | 639651 | PROSTUT05 | 729341X12 (LUNGNOT03), 839651CT1, 839651H1, and 839651X55 (PROSTUT05), 839651X60 (PROSTUT05) |
| 25 | 90 | 1253545 | LUNGFTETO3 | 1253545H1 and 12535414F6 (LUNGFTETO3), 18063337X13F1, and 1807400X11F1 (STINTNOT01), 2179882222F1 (STINTNOT01), 2552938F6 (LUNGNOT22), 2840018F6 (DRGQNOT01) |
| 26 | 91 | 1425691 | BEPTNON01 | 2727135H1 (OVARUTU05), 587126X29R1, 585598X17, and 587126F1 (UTRSNCNOT01), 1714529F6 (UCMCNOT02), 1381341F6 (BRAITUT08), 1273513F6 (TESTUTU02), 060265R1 (LUNGNOT01), 1459659F1 (COLANFETO2), 043139R1 (TBUDNOT01), 1425691H1 (BEPTNON01) |

Table 1 cont.

| Protein SEQ ID NO: | Nucleotide SEQ ID NO: | Clone ID | Library | Fragments |
|-----------------------|--------------------------|----------|-----------|--|
| 27 | 92 | 1484257 | CORPNOT02 | 400685H1, 404702F1, 404702R6, 404702X45C1, 404702X47C1, and 404702X86C1 (TMFR31T01), 1484257H1 (CORENOT02), 3395312H1 (UTRSNOT16) |
| 28 | 93 | 1732368 | BRSTUTU03 | 920006H1 (RATNOT02), 1732368F6 and 1732368H1 (BRSTUTU08), 2607269T6 (LUNGNOT07), 2654163T6 (THYMMOT04) |
| 29 | 94 | 1970914 | SKINB1T01 | 1549551R6 (PROSNOT06), 1575349H1 (KIDNOT03), 1870914H1 (SKINB1T01), 2365856T6 (ADRENOT07), SBAA00149F1 |
| 30 | 95 | 1910984 | CONNUTU01 | 855876X12 (BRATUT03), 123497H1 and 1241845H1 (LUNGNOT03), 1910984F6 and 1910984H1 (CONNUTU01), 3227655H1 (PROSBPT06) |
| 31 | 96 | 1943040 | HIPONOT01 | 824144R1 (PROSNOT06), 930281H1 (CERVNOT01), 1420545H1 (KIDNOT09), 178405H1 (BRAINOT10), 1941040H1 and 1943040R6 (HIPONOT01), 2122271H1 (BRSTNOT07), 2729723H1 (OVARTU04) |
| 32 | 97 | 2076520 | ISLTNTU01 | 419755R1 (BRSTNOT01), 934937R1 (KIDNOT05), 1460268H1 (COLNPFET02), 1599016H1 (BLADNOT03), 2076520H1 (ISLTNTU01), 2082255F6 (UTRSNOT08), 2184150F6 (SININOT01), 2884394F6 (SININOT02), 3726575H1 (BRSTNOT23), 3752456H1 (UTRSNOT18), 3764971H1 (BRSTNOT24), 4412005H1 (MONOTU01) |

Table I cont.

| Protein SEQ ID NO: | Nucleotide SEQ ID NO: | Clone ID | Library | Fragments |
|-----------------------|--------------------------|----------|------------|---|
| 33 | 98 | 2291241 | BRAINNO1 | 2291241CT1 and 2291241H1 (BRAINNO1), 2500586H1 (ADRETUT05) |
| 34 | 99 | 2329692 | COLNNOT11 | 158011F1 (ADENINB01), 1519452F1 (BLADTUT04), 1543875R1 (PROSTUT04), 2329692H1 2331530R6, and 2331530T6 (COLNNOT11), 2078791F6 (SPCANOT01) |
| 35 | 100 | 2474110 | THP1NOT03 | 863265H1 (BRAUTUT03), 1313444T1 (BLADTUT02), 1872631T6 and 1872669T6 (LEUKRNOT02), 2061219R6 (OVARNOT03), 2171863H1 (ENDOCNOT03), 2474110H1 (THPINOT03), 2690250H1 (LUNGNOT23), 2812791F6 (OVARNOT01) |
| 36 | 101 | 2495790 | ADRETUT05 | 1360339T1 (LUNGNOT12), 168792H1 (PROSTUT10), 1795321H1 (PROSTUT05), 190521F6 (OVARNOT07), 1907168F6 (OVARNOT07), 2495790H1 (ADRETUT05), 2557522F6 (BRAUTUT22) |
| 37 | 102 | 2661254 | ADRENNOT08 | 124185H1 (LUNGNOT03), 1545867R1 (PROSTUT04), 2325561H1 (OVARNOT02), 2661254H1 (ADRENNOT08), 2731457H1 (THP1A2S08) |
| 38 | 103 | 2674047 | KIDNNNOT19 | 489330H1 (INT2AGT01), 2059116R6 (OVARNOT03), 2059116T6 (OVARNOT03), 2674047F6 and 2674047H1 (KIDNNNOT19), 2805474H1 (BLADTUT08), 3076605H1 (BONEUT01), 308013776 (BRAINNOT01) |

Table 1 cont.

| Protein SEQ ID NO: | Nucleotide SEQ ID NO: | Clone ID | Library | Fragment(s) |
|-----------------------|--------------------------|----------|-----------|---|
| 39 | 104 | 2762174 | SRA1NS12 | 2573448T ³ (H1POA2701), 2762174H1 (BRAINOS12), SRA1A0504F1, SRA1A1683F1, SRA1A0674F1, SRA1A00837F1 |
| 40 | 105 | 2765991 | BRSTNOT12 | 082008R6 (HUVESTB01), 2127491T6 (KIDNOT05), 2765991F6 and 2765991H1 (BRSTNOT12), 314781H1 (PENCNOT05), SZAH0153F1, SZAH0135F1 |
| 41 | 106 | 2775157 | PANCNOT15 | 232541H1 (OVARNOT02), 2506571F6 and 2506671T6 (COCNOT01), 277515F6 and 277515H1 (PANCNOT15), 3376091F6 (PENGNOT01), 3412063H1 (BRSTTUS08) |
| 42 | 107 | 2918375 | THYMFBT03 | 227782P1 (PANCNOT01), 1225559H1 (CORNTUTU02), 151145H1 (LUNGNOT14), 2918375H1 (THYMFBT03) |
| 43 | 108 | 3149729 | ADRENONC4 | 605315F1 (BRSTTUT01), 3149729CM1 and 3149729H1 (ADRENONC4) |
| 44 | 109 | 3705995 | PENCNOT07 | 744201R1 (BRAITUT01), 2550322H1 (LUNGNOT06), 3705895H1 (PENGNOT07) |

Table 1 cont.

| Protein SEQ ID No.: | Nucleotide SEQ ID No.: | Clone ID | Library | Fragments |
|---------------------|------------------------|----------|-----------|--|
| 45 | 110 | 003256 | HMC1N0T01 | 003256R6, 003256T6, 003256X305F1, 003256X313F, 003256X315F1, and 009404H1 (HMC1N0T01), 43104R1 (TBLYN0T01), 413017R1 (BRTNOT01) |
| 46 | 111 | 156986 | THP1P1E02 | 01084R1 and 012909H1 (THP1P1B01), 156986H1 and 156986R1 (THP1P1B02), 1320255F1 (BIA0DNOT04), 15122255F1 (LUNGNOT04), 2061923T6 (OVARNOT03), 2388787F6 (THP1B2ZT01), 251766H2 (LIVRNOT04) |
| 47 | 112 | 319415 | EOSIHE02 | 285773H1, 285773R1, 319415H1, and 319415X19F1 (EOSIHBET02), 123145H1 (BRATNOT01), 1804042F6 (SINTNOT13), 187885SF6 (LEHNNOT03) |
| 48 | 113 | 635581 | NEUTM0T1 | 635581H1 (NEUTM0T01), 3045776F6 (HEANOT01) |
| 49 | 114 | 921803 | RATRNOT02 | 921803H1 (RATRNOT02), 1275128T6 (TESTT02), 1709959F6 (PROSNOT01), 241654TF6 (HNTTAZT01), 3016146H1 (MUSCNOT07), 3577250H1 (BRONNOT01) |
| 50 | 115 | 1250492 | LUNGFT03 | 691921X14F1 (LUNGFT02), 1250492F6, 1250492H1, and 1252265F2 (LUNGFT03), 136164F6 (LUNGNOT12), 309358F6 (LUNGNOT25), 404452H1 and 4048275H1 (LUNGNOT35), 414529SH1 (SINTT04) |
| 51 | 116 | 1427388 | SINTB0T01 | 126118H1 (SYNORAT05), 142783H1 and 142783T1 (SINTB0T01), 173769T6 (BRSTT008), 2551854H1 (LUNGNOT06) |
| 52 | 117 | 1448258 | PLACNOT02 | 144825H1 and 1448258R1 (PLACNOT02), 1484126F1 (CORPNOT02), 1856631F6 and 1856631X11F1 (PROSNOT18), 2690070F6 (LUNGNOT23), SAMAA00131F1 and SAWA00146F1 |

Table 1 cont.

| Protein SEQ ID NO: | Nucleotide SEQ ID NO: | Clone ID | Library | Fragments |
|-----------------------|--------------------------|----------|-----------|--|
| 53 | 118 | 1645941 | PROSTUT09 | 831680F6 (PROSTUT04), 1645941F6 and 1645941H1 (PROSTUT09), 174882F6 (STOMTUT02), 187081F6 (SKINBTT01), 187590F6 (LEUKNOT03), 231027F6 (NGANNNOT01) |
| 54 | 119 | 1646095 | PROSTUT09 | 1446095H1, 164605X3109F1, 164605X312F1 and 1646883F6 (PROSTUT09), SAHA02276F1 |
| 55 | 120 | 16886561 | PROSNOT15 | 1234124H1 (LUNGFEPO3), 1289156F6 (BASTNOT07), 142518F1 (BEP1) (NON01), 154715F1 (PROSTUT04), 166651H1 (PROSNOT15), 272310H1 (LUNGNOT10), 275215F6H1 (TRP12S08), 33580F6 (BRA1FET01), 350225F1 (ADREN011), 3857461H1 (LNDDNOT03), 5069547H1 (PANCNOT23) |
| 56 | 121 | 1821233 | GBLATUT01 | 010744H1 (THPINOB01), 1272043F1 (TESTUT02), 1419545F1 (KIDNNNOT09), 143377R1 (BEPINON01), 1482846F1 (CORPNOT02), 1821233H1 (GBBLATUT01), 1869022H1 (SERINBTT01) |
| 57 | 122 | 1877278 | LEUKNOT03 | 187114F6 (SKINBTT01), 1872728H1 (LEUKNOT03), 2097342T6 (BRA1UT02), 3124246F6 (LNDDNOT05), 3450007R6 (UTRSNOT03), 4894334H1 (LIVRTUT12), SAEB02108R1 |
| 58 | 123 | 1880692 | LEUKNOT03 | 1880692H1 (LEUKNOT03), SBA00446F1, SARA03727F1 |

Table 1 cont.

| Protein SEQ ID No: | Nucleotide SEQ ID No: | Clone ID | Library | Fragments |
|--------------------|-----------------------|----------|------------|---|
| 59 | 124 | 2280456 | PROSNON01 | 1557906F6 (BLADTUT04), 2280456H1 (PROSNON01), 2799446F6 |
| 60 | 125 | 2284580 | BRAINNON01 | 1215190T2 (BRSTTUT01), 1458188F1 (COLNENON03), 2284580H1 (BRAINNON01), 3519009H1 (LUNGNON03) |
| 61 | 126 | 2779172 | OVARTUT03 | 487548H1 and 487548R6 (HNT2ACT01), 1424684F1 (KIDNNON09), 217254F6 (ENDCNON03), 2672062F6 (ESGTTUT02), 277912F6 and 2779172H1 (INVARTUT03), 233502F6 (TMYMFET02), 3206879F6 (PENCNOT03) |
| 62 | 127 | 3279129 | STOMFET02 | 885282R6 and 885282T1 (PANCNOT05), 901139S1 (BRSTTUT03), 1655530F6 (PROSTTUT08), 1818669T6 (PROSISOT02), 2380664F6 (ISIUMNOT01), 28221229H1 (SININON04), 3279129H1 (STOMFET02), 3451425R6 (UTRSNON01) |
| 63 | 128 | 3340290 | SPLNNNOT10 | 102935H1 (ADRENON01), 1361133F6 (LUNGNOT12), 1674514H1 (BLADNOT05), 2271137H1 (PRFSNCN01), 2827770H1 (TLYMANOT03), 3340290H1 (SPINNOT10), 4556330H1 (NEBRAUNNOT01) |
| 64 | 129 | 3376404 | PENGNOT01 | 3376404H1, 3376404X0001, 3376404X310U1, and 3376404X323U1 (PENGNOT01), 3741323X302B1 (MENTNOT01) |
| 65 | 130 | 4173111 | SINTNOT21 | 1337315F6 (COLNNON13), 2486184F6 (CONUTTUT01), 4173111H1 (\$INTNOT21), 4750042H1 (SMCRUNNOT01) |

Table 2

| Protein SEQ ID NO: | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation sites | Signature Sequence | Identification | Analytical Methods |
|-----------------------|------------------------|--|-------------------------------------|-------------------------|--|-----------------------------|
| 1 | 155 | S9, S16, T25, S37, S56, S57, S61, S114, T152 | | G38-T73 | sigma-54 interaction protein | BLOCKS |
| 2 | 152 | S6, T83, S103, T121, S136 | | H99-R112 | LUPUS La protein | PRINTS |
| 3 | 304 | S30, S61, S94, T109, S132, S133, T183, T236, S277, S289 | N65, N294 | C228-C268 C231-I255 | zinc finger/RING finger protein | PFAM, BLOCKS |
| 4 | 178 | T8, S48, S102, Y121, T144 | | N119-P32 | histone H3 protein | PRINTS |
| 5 | 301 | T58, T70, T85, S148, T165, S256, T272, S81 | N191 | K21-F38 | filaggrin structural protein | PRINTS |
| 6 | 250 | S99, S126, S142, S155, T182 | | F203-V214 | maspin/breast tumor suppressor protein | PRINTS |
| 7 | 371 | T25, S46, S96, T123, S128, T144, S163, S167, S205, S221, T350 | N203, N222, N307, N348 | EQ165-Y185 K152-L192 | luman/leucine zipper/CRE protein | BLAST, BLOCKS, PRINTS |

Table 2 cont.

| Protein SEQ ID NO: | Amino Acid Residues | Potential Phosphorylation Sites | Potential glycosylation sites | Signature Sequence | Identification | Analytical Methods |
|-----------------------|------------------------|---|-------------------------------------|-----------------------|--|----------------------------|
| 8 | 148 | T35, S41, S92, S105 | N144 | | TSC-22 transcription factor | BLAST |
| 9 | 127 | T69 | N53 | M1-E16 | Ribosomal protein S6 | PFAM |
| 10 | 383 | S22, T34, S53, S140, T155, T183, S225, T263, S273, S300, S308, T369, S375 | N127 | Q7-K112 | PH-domain protein | PFAM |
| 11 | 254 | T57, S62, S92, S143, S146, T166, T176, S180, T187, S191, S194, T221 | | | Cyclin-dependent-k inase binding protein | BLAST |
| 12 | 305 | S65, T98, S146, S230, S246, S272 | N221 | G84-N271 | ribosomal protein L2 | PFAM, BLOCKS |
| 13 | 230 | T34, T49, S54, S122, T123, T150, S182, T209 | N86, N110, N199 | C155-C191 | Zinc finger/RING finger protein | PFAM, BLOCKS, NOTIFS |
| 14 | 292 | S2, T61, T89, T193, S223, S224, S225, S238, S288 | N47, N101, N166, N259 | A124-I145 | FOS transforming protein | PRINTS |

Table 2 cont.

| Protein SEQ ID NO: | Amino Acid Residues | Potential Phosphorylation Sites | Potential glycosylation sites | Signature Sequence | Identification | Analytical Methods |
|--------------------|---------------------|---|-------------------------------|------------------------|---------------------------------|---------------------|
| 15 | 232 | T58, S72, S127, S149, T154, S191, S199, T203, T204 | N56, N183, N187 | E39-F73 | tropomyosin | BLOCKS PRINTS |
| 16 | 376 | T5, T34, S53, T70, S81, T86, S105, S256, T287, T288, T310, S331, S364, S369, T365 | | Q97-C135 | RecA DNA repair protein | BLOCKS BLAST |
| 17 | 204 | T100, T118, T157, S187, S199 | | L179-H200 | annexin | PRINTS |
| 18 | 713 | S46, T64, T71, T95, S96, T139, T171, S260, S286, T345, S38, S485, T527, T541, Y567, Y593, S644, T656 | N110, N453, N460, N595 | L563-L576 L583-1596 | RSP-1 /Ras-signaling protein | BLAST, PRINTS |
| 19 | 360 | S22, T51, S69, T106, S133, S206, T232, S248 | | | Nucleolar protein Surf-6 | BLAST |
| 20 | 196 | S38 S69 T23 T30 S73 S183 S3 / T84 | N9 N51 | E76-L91 R35-K58 | Helix-loop-helix protein HES-1 | MOTIFS BLOCKS BLAST |

Table 2 cont.

| Protein Seq ID No: | Amino Acid Residues | Potential Phosphorylation Sites | Potential glycosylation sites | Signature Sequence | Identification | Analytical Methods |
|--------------------|---------------------|---|-------------------------------|---|------------------------------|---------------------------|
| 21 | 540 | T136 S14 S69 S189 T122 S111 T7 S66 S75 T119 S193 S197 S205 T285 S324 S328 S380 S425 | N240 N443 | H230-H252, C260-C264- H280, C283-H309, C316-H336, C344- H364, C372-H392, C400-H420, C428- H448, C456-H476, C184-H504, C512- H532 | zinc finger protein | MOTIFS BLAST PRINTS |
| 22 | 549 | S123 S32 S182 T319 T465 S161 T205 S208 S332 S392 S459 S534 | N167 N335 N422 | C214-H234, C242- H262, C270-H290, C298-H318, C326- H346, C354-H374, C382-H402, C410- H430, C438-H458, C466-H486, C494- H514, C522-H542 | zinc finger protein ZNF43 | MOTIFS BLAST PRINTS |
| 23 | 361 | S244 T254 S8 S58 S180 S113 T269 T233 S284 T26 S45 S174 T224 S314 | | C139-L163 C227-K263 | DNA binding protein | BLOCKS BLAST |
| 24 | 241 | S82 S62 S119 T147 Y111 | | C52-H75, C83- H105, C113-H113, C141-H161, C172- H193 | zinc finger protein PZF | MOTIFS PRINTS BLAST |

Table 2 cont.

| Protein Seq ID No: | Amino Acid Residues | Potential Phosphorylation Sites | Potential glycosylation sites | Signature Sequence | Identification | Analytical Methods |
|--------------------|---|---|-------------------------------------|---|----------------------------|--------------------|
| 25 576 | S90 T311 S56 T183 T195 S203 S316 T318 S347 S354 S412 S548 S37 S82 S281 T325 S343 S409 S414 S447 S466 T481 S502 S570 S323 | N42 N312 N339 N498 | C507-L513, L168- L189, E262-R278 | transcription factor | MOTIFS PRINTS BLOCKS BLAST | |
| 26 406 | S74 S197 T226 S247 T289 S328 S338 S353 S336 S334 T14 S199 S234 T388 | N245 N253 | G164-R175 | transcription factor | PRINTS BLAST | |
| 27 810 | S392 S113 S155 S185 S225 S262 S283 T288 S342 S443 T449 T665 T655 S728 T756 T801 T79 T190 S377 T438 Y397 | C315-H335, C343- H363, C371-H391, C399-H419, C427- H447, C455-H475, C483-H503, C511- H531, C539-H559, C567-H587, C595- H615, C623-H644, C726-H747 | zinc finger protein Miz-1 | MOTIFS PRINTS BLOCKS | | |
| 28 324 | S72 T189 S209 T223 S279 S302 S156 T182 S316 Y277 | N187 | C74-R85 | Hormone-binding transcription factor protein | PRINTS BLAST | |
| 29 292 | S242 T41 S136 S137 T176 T200 S205 S224 T52 S61 | N229 | G62-S69 | putative nucleotide-binding protein | MOTIFS PRINTS BLAST | |

Table 2 cont.

| Protein Seq ID No: | Amino Acid Residues | Potential Phosphorylation Sites | Potential glycosylation sites | Signature Sequence | Identification | Analytical Methods |
|--------------------|---------------------|--|-------------------------------|------------------------|--|---------------------------|
| 30 | 259 | T79 S99 S180 T20 S152 S241 | | C71-H92, C43-C71 | zinc finger protein | MOTIFS BLOCKS BLAST |
| 31 | 97 | S52 | | C15-L43 | DNA-binding protein | MOTIFS BLOCKS BLAST |
| 32 | 812 | T239 T76 S55 T56 T104 S126 S127 T155 S178 T249 S268 T269 S330 T394 S450 T484 S583 S652 S658 S795 S33 S235 T314 S343 T730 S804 | N45 N93 N165 N805 | E418-S450 | cell cycle protein | BLOCKS BLAST [®] |
| 33 | 392 | T22 S30 T43 S55 S108 T140 S156 S118 T320 S343 S120 S270 S311 | | R277 | TRAF family member-associated NF-κB activator TANK | BLAST |
| 34 | 60 | T49 T30 S50 | | I2-S55 | DNA-binding protein | BLOCKS BLAST |
| 35 | 209 | S21 S57 T93 | N67 | F160-N179 S151-G185 | cellular nucleic acid binding protein | PRINTS BLOCKS BLAST |
| 36 | 257 | T178 S187 S230 T249 | N65 | Y33-F44 S187-I205 | cell-cycle control protein Hst2p | PRINTS BLOCKS BLAST |

Table 2 cont.

| Protein Seq ID No: | Amino Acid Residues | Potential Phosphorylation Sites | Potential glycosylation sites | Signature Sequence | Identification | Analytical Methods |
|--------------------|---------------------|---|--|---|---------------------------------------|----------------------------|
| 37 | 138 | T106 T3 S27 S46 | | E108-Q124 | nucleic acid-binding protein | BLOCKS BLAST |
| 38 | 999 | T54 S64 S89 S126 S35 S41 S42 S451 T512 T162 T792 T858 S890 T97 T994 T205 S233 T274 T491 S525 S534 P577 T600 S610 S615 S634 S677 T951 S961 Y152 Y458 Y686 Y815 | N43 N532 M672 N749 N818 N943 T792 T858 S890 T97 T994 T205 S233 T274 T491 S525 S534 P577 T600 S610 S615 S634 S677 T951 S961 Y152 Y458 Y686 Y815 | L574-L595 L647-L668 | DNA-binding protein | MOTIFS BLAST |
| 39 | 377 | T142 T254 T48 T138 S222 S71 S74 S108 S114 T138 S222 S250 T332 T364 | | C130-H150, C158-H178, C186-H206, C214-H234, C242-H262, C270-H290, C296-H316, C324-H344, C332-H372 | zinc finger protein ZNF132 | MOTIFS PRINTS BLOCKS BLAST |
| 40 | 324 | S28 S314 S16 S81 S114 T225 T33 S44 T66 S303 S209 T229 | N47 | R26-S37 S77-L115 | transcription regulatory protein IRUB | PRINTS BLOCKS BLAST |
| 41 | 270 | S16 T123 T41 I199 S9 S55 S90 T128 T175 S194 S214 | N22 N109 N192 P250-Q263 | V218-L242 P250-Q263 | MOTIFS BLOCKS PRINTS | |

Table 2 cont.

| Protein Seq ID No: | Amino Acid Residues | Potential Phosphorylation Sites | Potential glycosylation sites | Signature Sequence | Identification | Analytical Methods |
|--------------------|---------------------|---|-------------------------------|---|-----------------------------------|---------------------------|
| 42 | 252 | T20 S88 S89 S101 T127 S218 T1121 S126 T152 | N33 N46 N216 N230 | Y-118, S88-F88, D159-S168 | cell-cycle control protein | PRINTS BLAST |
| 43 | 228 | T50 T107 T2 S42 S201 T31 S51 T52 T103 T107 T134 T143 T206 S210 T215 | N132 N141 N165 N197 | A38-S51, Q65-P100, S59-K89 | Transcriptional Repressor Protein | PRINTS BLOCKS BLAST |
| 44 | 117 | T93 T11 | | A86-E104 | CCAT-Binding Transcription factor | PRINTS BLAST |
| 45 | 252 | S83 T2 S57 T159 S250 Y102 | N197 | M1-S29 A85-K123 | Ribosomal protein | BLOCKS MOTIFS |
| 46 | 530 | T177 S234 S451 S519 T24 T238 | N217 N227 | TM Domains: Y147-A167 Y245-I262 L306-F325 L332-L351 S379-F399 L470-F489 | melibiose carrier protein | MOTIFS HMM |

Table 2 cont.

| SEQ ID NO: | Amino Acid Residues | Potential Phosphorylation Sites | Potential glycosylation sites | Signature Sequence | Identification | Analytical Methods |
|------------|---------------------|--|--|--|-----------------------------------|----------------------------------|
| 47 | 355 | S7 S21 T127 S213 T279 S134 T276 S315 S331 S334 Y193 T300 | N37 N192 N263 N268 N337 G171-G200 N234-I256 | I42-E69 W160-E187 G171-G200 N234-I256 | Mycelin P0 Protein | BLOCKS, PRINTS MOTIFS, IHM |
| 48 | 136 | T109 S130 T5 T69 T40 S121 | | | geminin | BLAST, MOTIFS |
| 49 | 235 | T138 T142 S180 S230 S111 S120 S137 T224 | N140 N198 | ATP/GTP binding: G9-T16 | PTB-associated splicing factor | BLAST MOTIFS |
| 50 | 70 | T2 S64 | | | ninjurin | BLAST MOTIFS |
| 51 | 169 | T128 T26 S96 | | | B locus M Beta chain 1 | BLAST, MOTIFS |
| 52 | 359 | S55 S78 T161 S245 T282 T350 T57 T130 T289 | N105 | E205-S242 E271-V294 | ribosomal protein S6 kinase 2 | BLAST, MOTIFS BLOCKS, PFAM |
| 53 | 545 | S235 T317 S47 S73 S114 S146 S184 S236 S241 S194 S538 S2 T84 S109 S124 T210 S231 S266 S310 T360 S379 S535 | N45 N139 N431 N478 N511 | K8-1106 A333-I362 | ribosomal protein | MOTIFS BLOCKS PRINTS |

Table 2 cont.

| SEQ ID NO: | Amino Acid Residues | Potential Phosphorylation Sites | Potential glycosylation sites | Signature Sequence | Identification | Analytical Methods |
|------------|---------------------|---|-------------------------------|--------------------|--|--------------------------------------|
| 54 | 99 | T90 T43 T76 | | | ORF E4 | BLAST, MOTIFS |
| 55 | 565 | S27 S56 S112 T112 T197 S319 T411 T429 S475 T66 S156 S303 T390 S463 Y549 | N2 N55 N165 | | Sec1 precursor | BLAST, MOTIFS |
| 56 | 197 | S65 T73 S102 S19 T60 Y61 S136 S147 | N20 | | Regulatory protein | BLAST, MOTIFS |
| 57 | 321 | S91 S119 T139 S283 S147 T300 Y238 | N103 N194 | | putative ras effector protein | BLAST, MOTIFS |
| 58 | 356 | T45 S55 S91 S95 T103 S114 T142 S168 T317 S340 S49 S58 T236 S258 S314 Y12 Y296 | N91 N312 | | weak similarity to S. cerevisiae intracellular transport protein | BLAST, MOTIFS |
| 59 | 299 | S273 T81 S116 S120 T122 S146 S66 S151 T210 S225 T268 | | | PI3 Kinase P85 Regulator | MOTIFS, PRINTS |
| 60 | 293 | T34 S218 S247 S290 S291 T240 S79 S145 T156 T199 S204 S283 | N152 | V47-V71 K86-F93 | RNA-binding protein | BLAST, MOTIFS, BLOCKS, PFAM |

Table 2 cont.

| Seq ID No: | Amino Acid Residues | Potential Phosphorylation Sites | Potential glycosylation sites | Signature Sequence | Identification | Analytical Methods | |
|------------|---------------------|---|---|---|------------------------------|--|--|
| 61 | 777 | S81 S128 S141 T230 S315 S342 S352 T519 S564 S576 S684 T699 T758 T205 S213 S236 S294 S397 T417 S470 S515 T560 S640 T746 | N228 N281 N319 N453 N481 N636 N682 | | Zinc finger helicase | BLAST, MOTIFS | |
| 62 | 97 | T83 | | C20-C28 | ferredoxin | MOTIFS | |
| 63 | 308 | S15 S81 T97 T102 S103 S135 S200 S238 S28 S131 T154 S171 S186 Y232 | N58 N78 N95 N198 N236 | | ubiquitin-conjugating enzyme | BLAST, MOTIFS | |
| 64 | 290 | S121 S165 S167 S248 S17 T188 T207 Y46 Y203 | N55 N79 | M1-A22 C60-C76 C225-C235 W249-T272 | prostasin | BLAST, MOTIFS, BLO CKS, PRINTS PFAM, HMM | |
| 65 | 198 | S7 S9 S56 T115 T34 T86 | N183 | | transcriptional regulator | BLAST MOTIFS | |

TABLE 3

| Nucleotide Seq ID No: | Tissue Expression (Fraction of Total) | Disease Class (Fraction of Total) | Vector |
|-----------------------|--|--|-------------|
| 66 | Nervous (0.256) Reproductive (0.209) | Cancer (0.442), Inflammation (0.279), Proliferative/Fetal (112%) | pBlueScript |
| 67 | Reproductive (0.274) Cardiovascular (0.194) | Cancer (0.484), Inflammation (0.145), Proliferative/Fetal (0.194) | pBlueScript |
| 68 | Reproductive (0.231) Cardiovascular (0.205) | Cancer (0.385), Inflammation (0.231), Proliferative/Fetal (0.205) | pBlueScript |
| 69 | Reproductive (0.215) Hematopoietic/Immune (0.190) | Cancer (0.397), Inflammation (0.314), Proliferative/Fetal (0.215) | pBlueScript |
| 70 | Reproductive (0.367) Cardiovascular (0.122) | Cancer (0.489), Inflammation (0.233), Proliferative/Fetal (0.189) | pBlueScript |
| 71 | Reproductive (0.292) Nervous (0.142) | Cancer (0.469), Inflammation (0.257), Proliferative/Fetal (0.177) | pSPORT1 |
| 72 | Reproductive (0.261) Nervous (0.157) | Cancer (0.493), Inflammation (0.194), Trauma (0.142) | pSPORT1 |
| 73 | Reproductive (0.343) Hematopoietic/Immune (0.200) | Cancer (0.457), Inflammation (0.257), Trauma (0.229) | pINCY |
| 74 | Reproductive (0.320) Nervous (0.160) | Cancer (0.507), Inflammation (0.187), Proliferative/Fetal (0.133) | pSPORT1 |
| 75 | Gastrointestinal (0.300) Nervous (0.250) | Cancer (0.400), Inflammation (0.300) | pINCY |
| 76 | Reproductive (0.262) Nervous (0.181) | Cancer (0.443), Inflammation (0.262), Proliferative/Fetal (0.230) | pINCY |
| 77 | Reproductive (0.283) Nervous (0.151) | Cancer (0.509), Inflammation (0.208), Trauma (0.132) | pINCY |

TABLE 3 cont.

| NUCLEOTIDE SEQ ID NO: | Tissue Expression (Fraction of Total) | Disease Class (Fraction of Total) | Vector |
|--------------------------|--|--|-------------|
| 78 | Cardiovascular (0.300) Nervous (0.200) | Cancer (0.450), Inflammation (0.200) | pBLUESCRIPT |
| 79 | Reproductive (0.270) Cardiovascular (0.150) | Cancer (0.440), Inflammation (0.180), Proliferative/Fetal (0.150) | pINCY |
| 80 | Reproductive (0.271) Cardiovascular (0.153) | Cancer (0.506), Inflammation (0.176), Proliferative/Fetal (0.188) | pSPORT1 |
| 81 | Hematopoietic/Immune (0.312) Reproductive (0.219) | Cancer (0.344), Inflammation (0.344), Proliferative/Fetal (0.281) | pINCY |
| 82 | Nervous (0.259) Hematopoietic/Immune (0.188) | Cancer (0.500), Inflammation (0.438), Proliferative/Fetal (0.188) | pINCY |
| 83 | Hematopoietic/Immune (0.276) Reproductive (0.276) | Cancer (0.552), Inflammation (0.310) | pINCY |
| 84 | Reproductive (0.309) Nervous (0.148) | Cancer (0.526), Inflammation (0.247), Proliferative/Fetal (0.134) | pINCY |
| 85 | Reproductive (0.315) Nervous (0.152) Cardiovascular (0.130) | Cancer (0.522) Fetal (0.174) | pBLUESCRIPT |
| 86 | Reproductive (0.545) Hematopoietic/Immune (0.182) Gastrointestinal (0.182) | Cancer (0.636) Fetal (0.273) Inflammation (0.182) | pBLUESCRIPT |
| 87 | Reproductive (0.218) Nervous (0.200) Hematopoietic/Immune (0.200) | Cancer (0.509) Inflammation (0.236) Fetal (0.164) | pSPORT1 |
| 88 | Nervous (0.297) Reproductive (0.185) Hematopoietic/Immune (0.148) | Cancer (0.407) Fetal (0.259) Inflammation (0.222) | pSPORT1 |

TABLE 3 cont.

| Nucleotide Seq ID No: | Tissue Expression (Fraction of Total) | | Disease Class (Fraction of Total) | | Vector |
|--------------------------|---|------------------|--|--------------|-------------|
| | Reproductive | Gastrointestinal | Cancer | Inflammation | |
| 89 | [0.339] Nervous (0.161) Gastrointestinal (0.145) Cardiovascular (0.145) | | Cancer (0.613) Fetal (0.145) Inflammation (0.129) | | pSPORT1 |
| 90 | Cardiovascular (0.278) Gastrointestinal (0.204) Reproductive (0.185) | | Cancer (0.519) Inflammation (0.204) Fetal (0.148) | | pINCY |
| 91 | Reproductive (0.228) Nervous (0.149) Gastrointestinal (0.146) | | Cancer (0.411) Inflammation (0.343) Fetal (0.240) | | pTT3 |
| 92 | Reproductive (0.240) Hematopoietic/Immune (0.160) Gastrointestinal (0.160) | | Cancer (0.460) Inflammation (0.260) Fetal (0.180) | | pINCY |
| 93 | Reproductive (0.333) Cardiovascular (0.200) Hematopoietic/Immune (0.133) | | Inflammation (0.533) Cancer (0.400) Fetal (0.133) | | pINCY |
| 94 | Reproductive (0.230) Gastrointestinal (0.164) Cardiovascular (0.115) Hematopoietic/Immune (0.115) | | Cancer (0.443) Inflammation (0.442) Fetal (0.197) | | pINCY |
| 95 | Reproductive (0.333) Cardiovascular (0.157) Gastrointestinal (0.167) | | Cancer (0.750) Inflammation (0.250) | | pINCY |
| 96 | Reproductive (0.369) Nervous (0.215) Hematopoietic/Immune (0.108) Gastrointestinal (0.108) | | Cancer (0.598) Inflammation (0.231) Fetal (0.108) | | pBLUESCRIPT |
| 97 | Reproductive (0.321) Gastrointestinal (0.179) Hematopoietic/Immune (0.161) | | Inflammation (0.411) Cancer (0.393) Fetal (0.161) | | pINCY |
| 98 | Reproductive (0.205) Nervous (0.192) Cardiovascular (0.164) | | Cancer (0.452) Inflammation (0.342) Fetal (0.178) | | pSPORT1 |

TABLE 3 cont.

| Nucleotide Seq ID NO: | Tissue Expression (Fraction of Total) | Disease Class (Fraction of Total) | Vector |
|-----------------------|--|--|------------------|
| 99 (0.115) | Gastrointestinal (0.423) Reproductive (0.281) Reproductive (0.234) Hematopoietic/Immune (0.141) | Cancer (0.385) Fetal (0.173) Cancer (0.375) Fetal (0.312) Inflammation (0.312) | pSPORT1 pINCY |
| 100 (0.141) | Reproductive (0.294) Nervous (0.196) Gastrointestinal (0.118) | Cancer (0.529) Fetal (0.255) | pINCY |
| 101 (0.118) | Reproductive (0.217) Nervous (0.163) Cardiovascular (0.141) | Cancer (0.435) Fetal (0.152) Cancer (0.526) Fetal (0.158) | pINCY |
| 102 (0.158) | Reproductive (0.263) Hematopoietic/Immune (0.158) Musculoskeletal (0.158) | Cancer (0.400) Fetal (0.263) | pINCY |
| 103 (0.158) | Reproductive (0.400) Reproductive (0.300) Reproductive (0.375) Cardiovascular (0.125) Urologic (0.125) | Cancer (0.400) Fetal (0.280) | pSPORT1 pINCY |
| 104 (0.100) | Gastrointestinal (0.400) Reproductive (0.400) Developmental (0.100) Hematopoietic/Immune (0.100) | Cancer (0.600) Fetal (0.200) Inflammation (0.200) | pINCY |
| 105 (0.139) | Reproductive (0.275) Gastrointestinal (0.152) Nervous (0.139) | Cancer (0.418) Fetal (0.165) | >pINCY |
| 106 (0.136) | Reproductive (0.364) Hematopoietic/Immune (0.182) Nervous (0.167) | Inflammation (0.409) Cancer (0.364) Fetal (0.136) | pSPORT1 pINCY |

TABLE 3 cont.

| Nucleotide Seq ID No: | Tissue Expression (Fraction of Total) | Disease Class (Fraction of Total) | Vector |
|-----------------------|--|---|-------------|
| 109 | Nervous (0.227) Reproductive (0.205) Cardiovascular (0.136) Urologic (0.136) Gastrointestinal (0.136) | Cancer (0.568) Fetal (0.136) | PINCY |
| 110 | Hematopoietic/Immune (0.400) Urologic (0.400) Reproductive (0.200) | Cell Proliferation (0.800) Inflammation (0.800) | pBluescript |
| 111 | Gastrointestinal (0.213) Hematopoietic/Immune (0.191) Nervous (0.191) | Cell proliferation (0.744) Inflammation (0.489) | pBluescript |
| 112 | Hematopoietic/Immune (0.405) Gastrointestinal (0.167) Cardiovascular (0.119) | Inflammation (0.619) Cell proliferation (0.381) | pBluescript |
| 113 | Hematopoietic/Immune (0.667) Cardiovascular (0.333) | Inflammation (1.000) | PSPORI1 |
| 114 | Cardiovascular (0.412) Nervous (0.235) Musculoskeletal (0.118) | Cell proliferation (0.765) Inflammation (0.353) | PSPORI1 |
| 115 | Cardiovascular (0.548) Reproductive (0.161) Developmental (0.129) | Cell proliferation (0.806) Inflammation (0.226) | PINCY |
| 116 | Reproductive (0.267) Cardiovascular (0.233) Hematopoietic/Immune (0.233) | Cell proliferation (0.467) Inflammation (0.500) | PINCY |
| 117 | Reproductive (0.400) Cardiovascular (0.167) Gastrointestinal (0.133) | Cell proliferation (0.600) Inflammation (0.267) | PINCY |
| 118 | Nervous (0.205) Reproductive (0.205) Other (0.154) | Cell proliferation (0.461) Inflammation (0.385) | PINCY |
| 119 | Reproductive (0.500) Nervous (0.167) Hematopoietic/Immune (0.167) | Cancer (0.500) Inflammation (0.167) Neurological (0.167) | PINCY |

TABLE 3 cont.

| Nucleotide Seq ID No: | Tissue Expression (Fraction of Total) | Disease Class (Fraction of Total) | Vector |
|-----------------------|---|--|---------|
| 120 | Reproductive (0.396) Cardiovascular (0.125) (0.125) Musculoskeletal (0.125) | Cell proliferation (0.750) Inflammation (0.209) | pINCY |
| 121 | Reproductive (0.248) Hematopoietic/Immune (0.194) Gastrointestinal (0.147) | Cell Proliferation (0.651) Inflammation (0.380) | pINCY |
| 122 | Nervous (0.264) Cardiovascular (0.132) Reproductive (0.132) | Cell Proliferation (0.547) Inflammation (0.396) | pINCY |
| 123 | Reproductive (0.242) Nervous (0.152) Urologic (0.152) | Cell proliferation (0.788) Inflammation (0.303) | pINCY |
| 124 | Nervous (0.333) Cardiovascular (0.167) (0.167) Hematopoietic/Immune (0.167) | Cell proliferation (0.667) Inflammation (0.500) | pSPORT1 |
| 125 | Reproductive (0.290) Cardiovascular (0.113) (0.161) Hematopoietic/Immune (0.113) | Cell proliferation (0.709) Inflammation (0.306) | pSPORT1 |
| 126 | Reproductive (0.360) Nervous (0.120) Urologic (0.110) | Cell proliferation (0.680) Inflammation (0.320) | pINCY |
| 127 | Reproductive (0.361) Gastrointestinal (0.145) (0.145) Nervous (0.145) | Cell proliferation (0.600) Inflammation (0.400) | pINCY |
| 128 | Cardiovascular (0.154) Gastrointestinal (0.154) Reproductive (0.154) | Cell proliferation (0.616) Inflammation (0.308) | pINCY |
| 129 | Urologic (1.000) | Cancer (1.000) | pINCY |
| 130 | Hematopoietic/Immune (0.214) Cardiovascular (0.143) Gastrointestinal (0.143) | Cell proliferation (0.428) Inflammation (0.357) | pINCY |

TABLE 4

| Protein SEQ ID NO: | Clone ID | Library | Library Comment |
|-----------------------|----------|-----------|--|
| 1 | 001106 | U937NOT01 | U937NOT01 Library was constructed at Stratagene (STR937207) using RNA isolated from U937 monocyte-like cell line (ATCC CRL1593) established from malignant cells obtained from the pleural effusion of a 37-year-old Caucasian male with diffuse histiocytic lymphoma. |
| 2 | 004586 | HMC1NOT01 | HMC1NOT01 Library was constructed using RNA isolated from HMC-1 human mast cell line derived from a 52-year-old female. Patient history included mast cell leukemia. Family history included atherosclerotic coronary artery disease, joint disorder involving multiple joints, cerebrovascular disease, and diabetes insipidus. |
| 3 | 052927 | FIBRNOT01 | FIBRNOT01 Library was constructed at Stratagene (STR937212) using RNA isolated from the WI38 lung fibroblast cell line derived from a 3-month-old Caucasian female fetus. |
| 4 | 082843 | HUVESPB01 | HUVESPB01 Library was constructed using RNA isolated from shear-stressed HIV-EC-C (ATCC CRL 1730), an endothelial cell line derived from the vein of a normal human umbilical. |
| 5 | 322349 | BOSIHET02 | BOSIHET02 Library was constructed using RNA isolated from peripheral blood cells spherosed from a 48-year-old Caucasian male. Patient history included hypereosinophilia. |
| 6 | 397653 | PITUNOT02 | PITUNOT02 Library was constructed using RNA (Clontech 6584-1) isolated from the pituitary gland of 87 male and female donors, 15 to 75 years old. |
| 7 | 673766 | CRBLNOT01 | CRBLNOT01 Library was constructed using RNA isolated from cerebellum tissue of a 69-year-old Caucasian male, who died from chronic obstructive pulmonary disease. Patient history included heart failure, myocardial infarction, hypertension, osteoarthritis, and tobacco use. |

TABLE 4 cont.

| Protein SEQ ID NO: | Clone ID | Library | Library Comment |
|--------------------|----------|-----------|--|
| 8 | 1504753 | BRAITU07 | BRAITU07 Library was constructed using RNA isolated from left frontal lobe tumor tissue removed from the brain of a 32-year-old Caucasian male during excision of a cerebral meningeal lesion. Pathology indicated low grade desmoplastic neuronal neoplasm. Family history included low atherosclerotic coronary artery disease. |
| 9 | 1760185 | PITUNOT03 | PITUNOT03 Library was constructed using RNA isolated from pituitary tissue of a 46-year-old Caucasian male who died from colon cancer. Patient history included arthritis and peptic ulcer disease. |
| 10 | 1805061 | SINTNOT13 | SINTNOT13 Library was constructed using RNA isolated from ileum tissue removed from a 25-year-old Asian female during a partial colectomy and temporary ileostomy. Pathology indicated moderately active chronic ulcerative colitis involving colonic mucosa from the distal margin to the ascending colon. Family history included hyperlipidemia, depressive disorder, malignant cervical neoplasm, and viral hepatitis A. |
| 11 | 1850120 | LUNGFE03 | LUNGFE03 Library was constructed using RNA isolated from lung tissue removed from a Caucasian female fetus who died at 20 weeks' gestation. The mother was given seven days of erythromycin treatment for bronchitis during the first trimester. |
| 12 | 1852290 | LUNGFE03 | |
| 13 | 1944530 | PITUNOT01 | PITUNOT01 Library was constructed using RNA (Clontech 6584-2) isolated from the normal pituitary glands of 18 male and female Caucasian donors, 15 to 70 years old, who died from trauma. |
| 14 | 2019742 | CORNNOT01 | CORNNOT01 Library was constructed using RNA isolated from mesentery fat tissue removed from a 71-year-old Caucasian male during a partial colectomy and permanent colostomy. Patient history included a cholecystectomy, viral hepatitis, and a hemangioma. Family history included atherosclerotic coronary artery disease, myocardial infarction, and extrinsic asthma. |

TABLE 4 cont.

| Protein SEQ ID NO: | Clone ID | Library | Library Comment |
|--------------------|----------|------------|---|
| 15 | 2056042 | BEPINOT01 | BEPINOT01 Library was constructed using RNA isolated from a bronchial epithelium (NBEB) primary cell line derived from a 54-year-old Caucasian male. |
| 16 | 2396682 | THP1AZT01 | THP1AZT01 Library was constructed using RNA isolated from THP-1 promonocyte cells treated for three days with 0.8 micromolar 5-aza-2'-deoxycytidine. THP-1 (ATCC TIB 202) is a human promonocyte line derived from peripheral blood of a 1-year-old Caucasian male with acute monocytic leukemia. |
| 17 | 2518733 | BRAIIUT01 | BRAIIUT01 Library was constructed using RNA isolated from brain tumor tissue removed from the midline frontal lobe of a 61-year-old Caucasian female during excision of a cerebral meningial lesion. Pathology indicated subfrontal meningothelial meningioma with no atypia. Patient history included depressive disorder; family history included cerebrovascular disease, senile dementia, hyperlipidemia, benign hypertension, atherosclerotic coronary artery disease, and congestive heart failure. |
| 18 | 2709055 | PONSAZT01 | PONSAZT01 Library was constructed using polyA RNA isolated from diseased dors tissue removed from the brain of a 74-year-old Caucasian male who died from Alzheimer's disease. |
| 19 | 2724537 | LUNGUT01 | LUNGUT01 Library was constructed using RNA isolated from lung tumor tissue removed from the left upper lobe of a 65-year-old Caucasian female during a segmental lung resection. Pathology indicated a metastatic grade 2 myxoid liposarcoma and metastatic grade 4 liposarcoma. Patient history included soft tissue cancer, breast cancer, and secondary lung cancer. Family history included benign hypertension. |
| 20 | 025918 | SPLENFET01 | SPLENFET01 Library was constructed at Stratagene using RNA isolated from a pool of fetal spleen tissue. 2x10 ⁶ primary clones were amplified to stabilize the library for long-term storage. Amplification may significantly skew sequence abundances. |

TABLE 4 cont.

| Protein SEQ ID NO: | Clone ID | Library | Library Comment |
|--------------------|----------|------------|---|
| 21 | 438283 | THYRNOT01 | THYRNOT01 Library was constructed using RNA isolated from thyroid tissue removed from a 64-year-old Caucasian female who died from congestive heart failure. |
| 22 | 619699 | PGANNOT01 | PGANNOT01 Library was constructed using RNA isolated from paraganglionic tumor tissue removed from the intra-abdominal region of a 46-year-old Caucasian male during exploratory laparotomy. Pathology indicated a benign paraganglioma and was associated with a grade 2 renal cell carcinoma, clear cell type, which did not penetrate the capsule. Surgical margins were negative for tumor. |
| 23 | 693452 | SYNORATO3 | SYNORATO3 Library was constructed using RNA isolated from the wrist synovial membrane tissue of a 56-year-old female with rheumatoid arthritis. |
| 24 | 839651 | PROSTUT05 | PROSTUT05 Library was constructed using RNA isolated from prostate tumor tissue removed from a 69-year-old Caucasian male during a radical prostatectomy. Pathology indicated adenocarcinoma (Gleason grade 3+4). Adenofibromatous hyperplasia was also present. Family history included congestive heart failure, multiple myeloma, hyperlipidemia, and rheumatoid arthritis. |
| 25 | 1253545 | LUNGFEV03 | LUNGFEV03 Library was constructed using RNA isolated from lung tissue removed from a Caucasian female fetus who died at 20 weeks' gestation. |
| 26 | 1425691 | BEPINON01 | BEPINON01 normalized bronchial epithelium library was constructed from 5.12 million independent clones from the BEPINON01 library. RNA was made from a bronchial epithelium primary cell line derived from a 54-year-old Caucasian male. The normalization and hybridization conditions were adapted from Sozzi et al., PNAS (1994) 91:9228, using a longer (24-hour) reannealing hybridization period. |
| 27 | 1484257 | CORDPNOT02 | CORDPNOT02 Library was constructed using RNA isolated from diseased corpus callosum tissue removed from the brain of a 74-year-old Caucasian male who died from Alzheimer's disease. |

TABLE 4 cont.

| Protein SEQ ID NO: | Clone ID | Library | Library Comment |
|--------------------|----------|-----------|--|
| 28 | 1732368 | BRSTTU08 | BRSTTU08 Library was constructed using RNA isolated from breast tumor tissue removed from a 41-year-old Caucasian female during unilateral extended simple mastectomy. Pathology indicated invasive nuclear grade 2-3 adenocarcinoma, ductal type, with 3 of 23 lymph nodes positive for metastatic disease. Greater than 50% of the tumor volume was <i>in situ</i> , both comedo and non-comedo types. Immunostains were positive for estrogen/progesterone receptors, and uninvolved tissue showed proliferative changes. The patient concurrently underwent a total abdominal hysterectomy. Patient history included valvuloplasty of mitral valve without replacement, rheumatic mitral insufficiency, and rheumatic heart disease. Family history included acute myocardial infarction, atherosclerotic coronary artery disease, and type II diabetes. |
| 29 | 1870914 | SKINBTO1 | SKINBTO1 Library was constructed using RNA isolated from diseased skin tissue of the left lower leg. Patient history included erythema nodosum of the left lower leg. |
| 30 | 1910984 | CONNUTU01 | CONNUTU01 Library was constructed using RNA isolated from a soft tissue tumor removed from the clival area of the skull of a 30-year-old Caucasian female. Pathology indicated chondroid chordoma with neoplastic cells reactive for keratin. |
| 31 | 1943040 | HIPONOT01 | HIPONOT01 Library was constructed using RNA isolated from the hippocampus tissue of a 72-year-old Caucasian female who died from an intracranial bleed. Patient history included nose cancer, hypertension, and arthritis. |
| 32 | 2076520 | ISLTNOT01 | ISLTNOT01 Library was constructed using RNA isolated from a pooled collection of pancreatic islet cells. |
| 33 | 2291241 | BRAINN001 | BRAINN001 Library was constructed and normalized from 4.88 million independent clones from the BRAINT03 library. RNA was made from brain tissue removed from a 26-year-old Caucasian male during cranioplasty and excision of a cerebral meningeal lesion. Pathology for the associated tumor tissue indicated a grade 4 oligoastrocytoma in the right fronto-parietal part of the brain. |

TABLE 4 cont.

| Protein SEQ ID NO: | Clone ID | Library | Library Comment |
|--------------------|----------|-----------|--|
| 34 | 2329692 | COLNNT011 | COLNNT011 The COLNNT011 library was constructed using RNA isolated from colon tissue removed from a 60-year-old Caucasian male during a left hemicolectomy. |
| 35 | 2474110 | THPINOT03 | THPINOT03 Library was constructed using RNA isolated from untreated THP-1 cells (ATCC TIB 202), a human promonocyte line derived from the peripheral blood of a 1-year-old Caucasian male with acute monocytic leukemia. |
| 36 | 2495790 | ADRETUT05 | ADRETUT05 Library was constructed using RNA isolated from adrenal tumor tissue removed from a 52-year-old Caucasian female during a unilateral adrenalectomy. Pathology indicated a pheochromocytoma. |
| 37 | 2661254 | ADRENOT08 | ADRENOT08 ADRENOT08 PINCY Library was constructed using RNA isolated from adrenal tissue removed from a 20-year-old Caucasian male, who died from head trauma. |
| 38 | 2674047 | KIDNNOT19 | KIDNNOT19 KIDNNOT19 PINCY Library was constructed using RNA isolated from kidney tissue removed a 65-year-old Caucasian male during an exploratory laparotomy and nephroureterectomy. Pathology for the associated tumor tissue indicated a grade 1 renal cell carcinoma within the upper pole of the left kidney. Patient history included malignant melanoma of the abdominal skin, benign neoplasm of colon, cerebrovascular disease, and umbilical hernia. Family history included cerebrovascular and cerebrovascular disease, and prostate cancer. |

TABLE 4 cont.

| Protein SEQ ID NO: | Clone ID | Library | Library Comment |
|--------------------|----------|------------|--|
| 39 | 2762174 | BRAINOS12 | BRAINOS12 pSOPR1 Library was constructed from 4.9 million clones from the BRAINOT03 library by subtraction of abundantly expressed clone pools. RNA was made from brain tissue removed from a 26-year-old Caucasian male during cranioplasty and excision of a cerebral meningial lesion. Pathology for the associated tumor tissue indicated a grade 4 oligoastrocytoma in the right fronto-parietal part of the brain. |
| 40 | 2765991 | BRSTNOT12 | BRSTNOT12 pINCY Library was constructed using RNA isolated from diseased breast tissue removed from a 32-year-old Caucasian female during a bilateral reduction mammoplasty. Pathology indicated nonproliferative fibrocystic disease. Family history included benign hypertension and atherosclerotic coronary artery disease. |
| 41 | 2775157 | PANCNOT15 | PANCNOT15 pINCY Library was constructed using RNA isolated from diseased pancreatic tissue removed from a 15-year-old Caucasian male during a exploratory laparotomy with distal pancreatectomy and total splenectomy. Pathology indicated islet cell hyperplasia. Family history included prostate cancer and cardiovascular disease. |
| 42 | 2918375 | THYMFET03 | THYMFET03 Library was constructed using RNA isolated from thymus tissue removed from a Caucasian male fetus. |
| 43 | 3149729 | ADRENON04 | ADRENON04 normalized adrenal gland library was constructed from 1.36 million independent clones from an adrenal tissue library. Starting RNA was made from adrenal gland tissue removed from a 20-year-old Caucasian male who died from head trauma. The library was normalized in two rounds using conditions adapted from Soares et al. (PNAS (1994) 91:9228-9232) and Bonaldo et al. (Genome Res (1996) 6: 791-806) and a significantly longer (48-hours/round) reannealing hybridization period. |
| 44 | 3705895 | PENCNNOT07 | PENCNNOT07 Library was constructed using RNA isolated from penis right corpora cavernosa tissue removed from a male. |

TABLE 4 cont.

| Protein SEQ ID No: | Clone ID | Library | Library Comment |
|--------------------|----------|-----------|---|
| 45 | 003256 | HMC1NOT01 | HMC1NOT01 library was constructed using RNA isolated from the HMC-1 human mast cell line derived from a 52-year-old female. Patient history included mast cell leukemia. |
| 46 | 156986 | THP1LB02 | THP1LB02 library was constructed by reamplification of THP1LB01, which was made using RNA isolated from THP-1 cells cultured for 48 hours with 100 ng/ml phorbol ester (PMA), followed by a 4-hour culture in media containing 1 ug/ml LPS. THP-1 (ATCC TIB 202) is a human promonocyte line derived from the peripheral blood of a 1-year-old male with acute monocyte leukemia [ref: int. J. Cancer (1980) 26:171]. |
| 47 | 319415 | EOSIET02 | EOSIET02 library was constructed using RNA isolated from peripheral blood cells apheresed from a 48-year-old Caucasian male. Patient history included hypereosinophilia. The cell population was determined to be greater than 77% eosinophils by Wright's staining. |
| 48 | 635561 | NEUTM01 | NEUTM01 library was constructed using RNA isolated from peripheral blood granulocytes collected by density gradient centrifugation through Ficoll-Hypaque. The cells were isolated from buffy coat units obtained from 20 unrelated male and female donors. Cells were cultured in 10 nM GM-CSF for 1 hour before washing and harvesting for total RNA preparation. |
| 49 | 921803 | RATRNOT02 | RATRNOT02 library was constructed using RNA isolated from the right atrium tissue of a 39-year-old Caucasian male, who died from a gunshot wound. |
| 50 | 1250492 | LUNGFE03 | LUNGFE03 library was constructed using RNA isolated from lung tissue removed from a Caucasian female fetus, who died at 20 weeks' gestation. |
| 51 | 1427838 | SINTBST01 | SINTBST01 library was constructed using RNA isolated from ileum tissue obtained from an 18-year-old Caucasian female during bowel anastomosis. Pathology indicated Crohn's disease of the ileum, involving 15 cm of the small bowel. Family history included cerebrovascular disease and atherosclerotic coronary artery disease. |

TABLE 4 cont.

| Protein SEQ ID No: | Clone ID | Library | Library Comment |
|--------------------|----------|-----------|---|
| 52 | 1448238 | PLACNOT02 | PLACNOT02 library was constructed using RNA isolated from the placental tissue of a Hispanic female fetus, who was prematurely delivered at 21 weeks' gestation. Serologies of the mother's blood were positive for CMV (cytomegalovirus). |
| 53 | 1645941 | PROSTUT09 | PROSTUT09 library was constructed using RNA isolated from prostate tumor tissue removed from a 66-year-old Caucasian male during a radical prostatectomy, radical cystectomy, and urinary diversion. Pathology indicated grade 3 transitional cell carcinoma. The patient presented with prostatic inflammatory disease. Patient history included lung neoplasm and benign hypertension. Family history included a malignant breast neoplasm, tuberculosis, cerebrovascular disease, atherosclerotic coronary artery disease and lung cancer. |
| 54 | 1646005 | PROSTUT09 | PROSTUT09 library was constructed using RNA isolated from prostate tumor tissue removed from a 66-year-old Caucasian male during a radical prostatectomy, radical cystectomy, and urinary diversion. Pathology indicated grade 3 transitional cell carcinoma. The patient presented with prostatic inflammatory disease. Patient history included lung neoplasm and benign hypertension. Family history included a malignant breast neoplasm, tuberculosis, cerebrovascular disease, atherosclerotic coronary artery disease and lung cancer. |
| 55 | 1686561 | PROSN015 | PROSN015 library was constructed using RNA isolated from diseased prostate tissue removed from a 66-year-old Caucasian male during radical prostatectomy and regional lymph node excision. Pathology indicated adenofibromatous hyperplasia. Pathology for the associated tumor tissue indicated an adenocarcinoma (Gleason grade 2+3). The patient presented with elevated prostate specific antigen (PSA). Family history included prostate cancer, secondary bone cancer, and benign hypertension. |

TABLE 4 cont.

| Protein SEQ ID No: | Clone ID | Library | Library Comment |
|--------------------|----------|----------|--|
| 56 | 1821233 | GBLATU01 | The GBLATU01 library was constructed using RNA isolated from gallbladder tumor tissue removed from a 78-year-old Caucasian female during a cholecystectomy. Pathology indicated invasive grade 2 squamous cell carcinoma forming a mass in the gallbladder. Patient history included diverticulitis of the colon, palpitations, benign hypertension, and hyperlipidemia. Family history included a cholecystectomy, atherosclerotic coronary artery disease, atherosclerotic coronary artery disease, hyperlipidemia, and benign hypertension. |
| 57 | 1877278 | LEURN03 | The LEURN03 library was constructed using RNA isolated from white blood cells of a 27-year-old female with blood type A-. The donor tested negative for cytomegalovirus (CMV). |
| 58 | 1880632 | LEUKN03 | The LEUKN03 library was constructed using RNA isolated from white blood cells of a 27-year-old female with blood type A+. The donor tested negative for cytomegalovirus (CMV). |
| 59 | 2280456 | PROSN01 | The PROSN01 library was constructed and normalized from 4.4 Million independent clones from the PROSN01 library. RNA was made from prostate tissue removed from 28-year-old Caucasian male who died from a self-inflicted gunshot wound. The normalization and hybridization conditions were adapted from Soares, M.B. et al. (1994) Proc. Natl. Acad. Sci. USA 91:9228-9232 using a longer (19 hour) reannealing hybridization period. |
| 60 | 2284580 | BRAIN01 | The BRAIN01 library was constructed and normalized from 4.88 million independent clones from the BRAIN03 library. RNA was made from brain tissue removed from a 26-year-old Caucasian male during cranioplasty and excision of a cerebral meningeal lesion. Pathology for the associated tumor tissue indicated a grade 4 oligoastrocytoma in the right fronto-parietal part of the brain. |

TABLE 4 cont.

| Protein SEQ ID NO: | Clone ID | Library | Library Comment |
|--------------------|----------|------------|---|
| 61 | 2779172 | OVARUTU03 | OVARUTU03 library was constructed using RNA isolated from ovarian tumor tissue removed from the left ovary of a 52-year-old mixed ethnicity female during a total abdominal hysterectomy, bilateral salpingo-oophorectomy, peritoneal and lymphatic structure biopsy, regional lymph node excision, and peritoneal tissue destruction. Pathology indicated an invasive grade 3 (of 4) seroneoplastic carcinoma forming a mass in the left ovary. Patient history included breast cancer, chronic peptic ulcer, and joint pain. Family history included colon cancer, cerebrovascular disease, breast cancer, type II diabetes, esophagus cancer, and depressive disorder. |
| 62 | 3279329 | STOMFET02 | STOMFET02 library was constructed using RNA isolated from stomach tissue removed from a Hispanic male fetus, who died at 18 weeks' gestation. |
| 63 | 3340290 | SPLANN0110 | SPLANN0110 library was constructed using RNA isolated from spleen tissue removed from a 59-year-old Caucasian male during a total splenectomy and exploratory laparotomy. Pathology for the spleen indicated splenomegaly with congestion. The lymph nodes showed reactive follicular hyperplasia. The liver showed mild, nonspecific steatosis. The patient presented with abdominal pain, bloating of the abdomen, low-grade fever, and diaphoresis. Family history included myocardial infarction, arteriosclerotic cardiovascular disease, primary tuberculous infection, cerebrovascular disease and lymphoma. |
| 64 | 3376404 | PENGNOT01 | PENGNOT01 library was constructed using RNA isolated from glans tissue removed from the penis of a 3-year-old Black male. Pathology for the associated tumor tissue indicated invasive grade 4 urothelial carcinoma forming a soft tissue scrotal mass that invaded the cavernous body of the penis and encased both testicles. |
| 65 | 4173111 | SINTNOT21 | SINTNOT21 library was constructed using RNA isolated from small intestine tissue obtained from a 8-year-old Black male, who died from anoxia. Serology was negative. |

Table 5

| Program | Description | Reference | Parameter Threshold |
|-------------------|---|--|---|
| ABI FACTURA | A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences. | Perkin-Elmer Applied Biosystems, Foster City, CA. | Mismatch <50% |
| ABI/PARACEL FDF | A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences. | Perkin-Elmer Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA. | |
| ABI AutoAssembler | A program that assembles nucleic acid sequences. | Perkin-Elmer Applied Biosystems, Foster City, CA. | |
| BLAST | A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastt, tblastn, and tblastx. | Altschul, S.F. et al. (1990) <i>J. Mol. Biol.</i> 215:403-410; Altschul, S.F. et al. (1997) <i>Nucleic Acids Res.</i> 25:3389-3402. | ESTs: Probability value = 1.0E-8 or less Full Length sequences: Probability value = 1.0E-10 or less |
| | | Pearson, W.R. and D.J. Lipman (1988) <i>Proc. Natl. Acad. Sci.</i> 85:2444-2448; Pearson, W.R. (1990) <i>Methods Enzymol.</i> 183: 63-98; and Smith, T.F. and M.S. Waterman (1981) <i>Adv. Appl. Math.</i> 2:482-489. | ESTs; fasta E value = 1.0E-6 Assembled ESTs; fasta identity = 95% or greater and Match length=20 bases or greater; fastx E value=1.0E-3 or less Full Length sequences; fastx score=100 or greater |
| FASTA | A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises at least five functions: fasta, fastx, fasts, and search. | Pearson, W.R. and D.J. Lipman (1988) <i>Proc. Natl. Acad. Sci.</i> 85:2444-2448; Pearson, W.R. (1990) <i>Methods Enzymol.</i> 183: 63-98; and Smith, T.F. and M.S. Waterman (1981) <i>Adv. Appl. Math.</i> 2:482-489. | Score=100 or greater. Ratio of Score/Strength = 0.75 or larger; and, if applicable, Probability value = 1.0E-3 or less |
| | | Henikoff, S and J.G. Henikoff, <i>Nucl. Acid Res.</i> , 19:6365-72; 1991. J.G. Henikoff and S. Henikoff (1996) <i>Methods Enzymol.</i> 266:88-105; and Attwood, T.K. et al. (1997) <i>J. Chem. Inf. Comput. Sci.</i> 37:417-424. | Score=10-50 bits for PFAM hits, depending on individual protein families |
| BLIMPS | A BLOCKs IMProved Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions. | Krogh, A. et al. (1994) <i>J. Mol. Biol.</i> 235:1501-1531; Sonnhammer, E.L.L. et al. (1998) <i>Nucleic Acids Res.</i> 26:320-322. | Score=10-50 bits for PFAM hits, depending on individual protein families |
| HMMER | An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM. | | |

Table 5 cont.

| Program | Description | Reference | Parameter Threshold |
|-------------|--|--|---|
| PrositeScan | An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite. | Gribskov, M. et al. (1988) CABIOS 4:61-66, Gribskov, M. et al. (1989) Methods Enzymol. 183:146-159; Barroch, A. et al. (1997) Nucleic Acids Res. 25: 217-221. | Score=4.0 or greater |
| Phred | A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability. | Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194. | Score=120 or greater, Match length= 36 or greater |
| Phrap | A Phil's Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences. | Smith, J.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M.S. Waterman (1981) J. Mol. Biol. 147:195-197, and Green, P., University of Washington, Seattle, WA. | Score=120 or greater, Match length= 36 or greater |
| Consed | A graphical tool for viewing and editing Phrap assemblies | Gordon, D. et al. (1998) Genome Res. 8:195-202. | Score=5 or greater |
| SPSScan | A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides. | Nielson, H. et al. (1997) Protein Engineering 10(1-6); Claverie, J.M. and S. Audic (1997) CABIOS 12: 431-439. | Score=5 or greater |
| Modis | A program that searches amino acid sequences for patterns that matched those defined in Prosite. | Barroch et al. (1998) Wisconsin Package Program Manual, version 9, page M51.59, Genetics Computer Group, Madison, WI. | Score=4.0 or greater |

What is claimed is:

1. A substantially purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-65, and fragments thereof.
- 5 2. A substantially purified variant having at least 90% amino acid sequence identity to the amino acid sequence of claim 1.
3. An isolated and purified polynucleotide encoding the polypeptide of claim 1.
4. An isolated and purified polynucleotide variant having at least 70% polynucleotide sequence identity to the polynucleotide of claim 3.
- 10 5. An isolated and purified polynucleotide which hybridizes under stringent conditions to the polynucleotide of claim 3.
6. An isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide of claim 3.
7. A method for detecting a polynucleotide, the method comprising the steps of:
 - 15 (a) hybridizing the polynucleotide of claim 6 to at least one nucleic acid in a sample, thereby forming a hybridization complex; and
 - (b) detecting the hybridization complex, wherein the presence of the hybridization complex correlates with the presence of the polynucleotide in the sample.
8. The method of claim 7 further comprising amplifying the polynucleotide prior to hybridization.
- 20 9. An isolated and purified polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:66-130, and fragments thereof.
10. An isolated and purified polynucleotide variant having at least 70% polynucleotide sequence identity to the polynucleotide of claim 9.
- 25 11. An isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide of claim 9.
12. An expression vector comprising at least a fragment of the polynucleotide of claim 3.
13. A host cell comprising the expression vector of claim 12.
- 30 14. A method for producing a polypeptide, the method comprising the steps of:
 - a) culturing the host cell of claim 13 under conditions suitable for the expression of the polypeptide; and
 - b) recovering the polypeptide from the host cell culture.
15. A pharmaceutical composition comprising the polypeptide of claim 1 in conjunction with a suitable pharmaceutical carrier.

16. A purified antibody which specifically binds to the polypeptide of claim 1.
 17. A purified agonist of the polypeptide of claim 1.
 18. A purified antagonist of the polypeptide of claim 1.
 19. A method for treating or preventing a disorder associated with decreased expression or activity of HTRM, the method comprising administering to a subject in need of such treatment an effective amount of the pharmaceutical composition of claim 15.
- 5 20. A method for treating or preventing a disorder associated with increased expression or activity of HTRM, the method comprising administering to a subject in need of such treatment an effective amount of the antagonist of claim 18.

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

| | | | |
|--|---------------------------|--|--|
| (51) International Patent Classification ⁶ : | | (11) International Publication Number: | WO 99/57144 |
| C07K 14/47 | | A2 | (43) International Publication Date: 11 November 1999 (11.11.99) |
| (21) International Application Number: | | PCT/US99/09935 | |
| (22) International Filing Date: | | 4 May 1999 (04.05.99) | |
| (30) Priority Data: | | | |
| 60/084,254 | 5 May 1998 (05.05.98) | US | CA 94040 (US); BANDMAN, Olga [US/US]; 366 Anna Avenue, Mountain View, CA 94043 (US), LAL, Preeti [IN/US]; 2382 Lass Drive, Santa Clara, CA 95054 (US). |
| 60/095,827 | 7 August 1998 (07.08.98) | US | YUE, Henry [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US); REDDY, Roopa [IN/US]; 1233 W. McKinley Drive, Sunnyvale, CA 94086 (US); TANG, Y., Tom [CN/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). |
| 60/102,745 | 2 October 1998 (02.10.98) | US | GERSTIN, Edward, H. [US/US]; 1408 38th Avenue, San Francisco, CA 94122 (US); PATTERSON, Chandra [US/US]; 490 Sherwood Way #1, Menlo Park, CA 94025 (US); BAUGHN, Mariah, R. [US/US]; 1424 Santiago Road, San Leandro, CA 94577 (US); AZIMZAI, Yalda [US/US]; 2045 Rock Springs Drive, Hayward, CA 94547 (US); LU, Dyung, Aina, M. [US/US]; 55 Park Belmont Place, San Jose, CA 95136 (US). |
| (63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Applications | | (74) Agents: BILLINGS, Lucy, J. et al.; Incyte Pharmaceuticals, Inc., 3174 Porter Drive, Palo Alto, CA 94304 (US). | |
| US | 60/084,254 (CIP) | | |
| Filed on | 5 May 1998 (05.05.98) | | |
| US | 60/095,827 (CIP) | | |
| Filed on | 7 August 1998 (07.08.98) | | |
| US | 60/102,745 (CIP) | | |
| Filed on | 2 October 1998 (02.10.98) | | |
| (71) Applicant (for all designated States except US): INCYTE PHARMACEUTICALS, INC. [US/US]; 3174 Porter Drive, Palo Alto, CA 94304 (US). | | (81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). | |
| (72) Inventors; and | | Published | |
| (75) Inventors/Applicants (for US only): HILLMAN, Jennifer, L. [US/US]; 230 Monroe Drive #12, Mountain View, | | Without international search report and to be republished upon receipt of that report. | |

(54) Title: HUMAN TRANSCRIPTIONAL REGULATOR MOLECULES

(57) Abstract

The invention provides human transcriptional regulator molecules (HTRM) and polynucleotides which identify and encode HTRM. The invention also provides expression vectors, host cells, antibodies, agonists and antagonists. The invention also provides methods for diagnosing, treating or preventing disorders associated with expression of HTRM.

Docket No.: PF-0509 USN

**DECLARATION AND POWER OF ATTORNEY FOR
UNITED STATES PATENT APPLICATION**

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name, and I believe that I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if more than one name is listed below) of the subject matter which is claimed and for which a United States patent is sought on the invention entitled

HUMAN TRANSCRIPTIONAL REGULATOR MOLECULES

the specification of which:

/ is attached hereto.

/ X / was filed on To Be Assigned as application Serial No. 09/674,743 and if this box contains an X /, was amended on _____.

/ X / was filed as Patent Cooperation Treaty international application No. PCT/US99/09935 on May 4, 1999, if this box contains an X /, was amended on under Patent Cooperation Treaty Article 19 on _____ 2001, and if this box contains an X /, was amended on _____.

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge my duty to disclose information which is material to the examination of this application in accordance with Title 37, Code of Federal Regulations, §1.56(a).

I hereby claim the benefit under Title 35, United States Code, §119 or §365(a)-(b) of any foreign application(s) for patent or inventor's certificate indicated below and of any Patent Cooperation Treaty international applications(s) designating at least one country other than the United States indicated below and have also identified below any foreign application(s) for patent or inventor's certificate and Patent Cooperation Treaty international application(s) designating at least one country other than the United States for the same subject matter and having a filing date before that of the application for said subject matter the priority of which is claimed:

Docket No.: PF-0509 USN

| Country | Number | Filing Date | Priority Claimed |
|---------|--------|-------------|------------------|
| | | | / Yes / No |
| | | | / Yes / No |

I hereby claim the benefit under Title 35, United States Code, §119(e) of any United States provisional application(s) listed below.

| Application Serial No. | Filed | Status (Pending, Abandoned, Patented) |
|---------------------------|-----------------|--|
| 60/084,254 | May 4, 1998 | Expired |
| 60/095,827 | August 7, 1998 | Expired |
| 60/102,745 | October 2, 1998 | Expired |

I hereby claim the benefit under Title 35, United States Code, §120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in said prior application(s) in the manner required by the first paragraph of Title 35, United States Code §112, I acknowledge my duty to disclose material information as defined in Title 37 Code of Federal Regulations, §1.56(a) which occurred between the filing date(s) of the prior application(s) and the national or Patent Cooperation Treaty international filing date of this application:

| Application Serial No. | Filed | Status (Pending, Abandoned, Patented) |
|---------------------------|-------|--|
|---------------------------|-------|--|

I hereby appoint the following:

| | |
|----------------------|-----------------|
| Lucy J. Billings | Reg. No. 36,749 |
| Michael C. Cerrone | Reg. No. 39,132 |
| Diana Hamlet-Cox | Reg. No. 33,302 |
| Richard C. Ekstrom | Reg. No. 37,027 |
| Barrie D. Greene | Reg. No. 46,740 |
| Lynn E. Murry | Reg. No. 42,918 |
| Shirley A. Recipon | Reg. No. 47,016 |
| Susan K. Sather | Reg. No. 44,316 |
| Michelle M. Stempien | Reg. No. 41,327 |
| David G. Streeter | Reg. No. 43,168 |
| P. Ben Wang | Reg. No. 41,420 |

respectively and individually, as my patent attorneys and/or agents, with full power of substitution and revocation, to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith. Please address all communications to:

Docket No.: PF-0509 USN

**LEGAL DEPARTMENT
INCYTE GENOMICS, INC.
3160 PORTER DRIVE, PALO ALTO, CA 94304**

TEL: 650-855-0555 FAX: 650-849-8886 or 650-845-4166

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

First Joint Inventor:

1-0U

Full name: Jennifer L. Hillman
Signature: Jenifer L. Hillman
Date: October 5, 2001
Citizenship: United States
Residence: Mountain View, California CA
P.O. Address: 230 Monroe Drive, #17
Mountain View, California 94040

Second Joint Inventor:

2-00

Full name: Olga Bandman
Signature: Olga Bandman
Date: 26 September, 2001
Citizenship: United States
Residence: Mountain View, California CA
P.O. Address: 366 Anna Avenue
Mountain View, California 94043

Docket No.: PF-0509 USN

Third Joint Inventor: Full name: Preeti G. Lal3-00 Signature: Preeti LalDate: September 10, 2001Citizenship: IndiaResidence: Santa Clara, California CAP.O. Address: P.O. Box 5142
Santa Clara, California 95056**Fourth Joint Inventor:** Full name: Henry Yue4-00 Signature: Henry YueDate: September 24, 2001Citizenship: United StatesResidence: Sunnyvale, California CAP.O. Address: 826 Lois Avenue
Sunnyvale, California 94087**Fifth Joint Inventor:** Full name: Roopa Reddy5-00 Signature: M ReddyDate: September 17th, 2001Citizenship: IndiaResidence: Sunnyvale, California CAP.O. Address: 1233 W. McKinley Avenue, #3
Sunnyvale, California 94086

Docket No.: PF-0509 USN

Sixth Joint Inventor:

6-00

Full name: Y. Tom Tang
Signature: U. Tang
Date: Sept. 10, 2001
Citizenship: United States
Residence: San Jose, California CA
P.O. Address: 4230 Ranwick Court
San Jose, California 95118

Seventh Joint Inventor:

7-00

Full name: Edward H. Gerstlin
Signature: E. Gerstlin
Date: September 11, 2001
Citizenship: United States
Residence: San Francisco, California CA
P.O. Address: 1408 38th Avenue
San Francisco, California 94122

Eighth Joint Inventor:

8-00

Full name: Chandra Arvizu
Signature: Chandra Arvizu
Date: September 10, 2001
Citizenship: United States
Residence: Menlo Park, California CA
P.O. Address: 490 Sherwood Way, #1
Menlo Park, California 94025

Docket No.: PF-0509 USN

Ninth Joint Inventor:

9-00

Full name:

Mariah R. Baughn

Signature:

Mariah R. Baughn

Date:

September 24, 2001

Citizenship:

United States

Residence:

San Leandro, California

CA

P.O. Address:

14244 Santiago Road

San Leandro, California 94577

Tenth Joint Inventor:

10-00

Full name:

Yalda Azimzai

Signature:

Yalda Azimzai

Date:

September 13, 2001

Citizenship:

United States

Residence:

Castro Valley, California

CA

P.O. Address:

5518 Boulder Canyon Drive

Castro Valley, California 94552

Eleventh Joint Inventor:

11-00

Full name:

Dyung Aina M. Lu

Signature:

Aine Lu

Date:

Sept 7, 2001

Citizenship:

United States

Residence:

San Jose, California

CA

P.O. Address:

233 Coy Drive

San Jose, California 95123